CANADIAN THESES MADURO, MORRIS FORD ISBN 0-612-29069-7
CHARACTERIZATION OF THE NQ-29069
THESES CANADIENNES UNC-119 GENE OF CAENORHABDITIS ... 1 of 4

NOTE TO USERS

The original manuscript received by UMI contains pages with indistinct and/or slanted print. Pages were microfilmed as received.

This reproduction is the best copy available

UMI

UNIVERSITY OF ALBERTA

Characterization of the *unc-119* gene of *Caenorhabditis elegans*

BY

Morris Ford Maduro ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial
fulfillment of the requirements for the degree of Doctor of Philosophy
in
Molecular Biology and Genetics

DEPARTMENT OF BIOLOGICAL SCIENCES

Edmonton, Alberta
Spring, 1998

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-29069-7

Canadä

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "CHARACTERIZATION OF THE *unc-119* GENE OF *Caenorhabditis elegans*" submitted by Morris Ford Maduro in partial fulfillment for the degree of Doctor of Philosophy in Molecular Biology and Genetics.

Dave
Dr. D. Pilgrim

Dr. J. Goldberg

H. McDermid
Dr. H. McDermid

Dr. M. Russell

J. Culotti
Dr. J. Culotti

Dr. M. Walter

Jan 27/98 DP

Abstract

This thesis describes the cloning and characterization of the *unc-119* gene from the nematode *Caenorhabditis elegans.* Mutants display uncoordinated locomotion, constitutive pharyngeal pumping, and have an imipramine-resistant egg laying defect. In addition, animals have structural defects in their sensory neurons and are unable to enter the dauer state. As muscle ultrastructure is normal, the phenotype is consistent with a pleiotropic nervous system defect.

Three new *unc-119* alleles, with similar phenotypes, were created with the mutagen EMS. A long-range restriction map of the *unc-119* region was developed, making use of a strain which is homozygous for a chromosomal rearrangement in the region. The *unc-119* region was subcloned from yeast artificial chromosome (YAC) and cosmid clones; isolation was confirmed by microinjection rescue.

The *unc-119* gene expresses a single mRNA of size ~980 bases, made up of five exons, with a single open reading frame (ORF) of 219 amino acids. The original allele, *e2498*, is a transposon insertion, while the three EMS alleles are point mutations that are predicted nulls. The *C. briggsae* homologue of *unc-119* was also cloned and predicted to encode a 217 a.a. polypeptide with 90% identity to the *C. elegans* ORF. Reporter gene fusions to *lacZ* or to the Green Fluorescent Protein (GFP) suggest that *unc-119* expression begins in the early embryo and continues through adulthood in most, if not all, neurons.

A closer examination showed that *unc-119* mutants have defasciculated ventral nerve cords and abnormally branched commissural axons. This suggests that UNC-119 acts during growth cone elongation, perhaps in manifesting the intracellular cytoskeletal rearrangements that accompany the interpretation of external guidance cues.

The predicted UNC-119 protein is very similar to that encoded by HRG4, a vertebrate retina-specific transcript. An *unc-119*::HRG4 transgene fully rescues the *unc-119* defects. Functional UNC-119::GFP and HRG4::GFP fusions localize to axons, consistent with a cytoplasmic function. A new homologue, DmUNC119, has been identified from the fruit fly *Drosophila*, which can also substitute for *C. elegans unc-119*. These results demonstrate that *unc-119* is an evolutionarily conserved gene.

Acknowledgements

I would like to thank my supervisor, David Pilgrim, for his help, advice and encouragement. He has exceeded my expectations of a thesis advisor and was a joy to work with. I also extend gratitude to the other members of the Pilgrim lab over the years for making the lab an enjoyable place, especially Petra Jäckle-Baldwin, David Hansen, Paul Stothard, Wanyuan Ao, Angela Manning, Wayne Materi, Jessica Smith, and Janelle Jakubowski.

I also thank Ariel Finkielsztein and Jack von Borstel for many non-science discussions.

I owe a debt of gratitude to the members of my supervisory committee, Mike Russell, Heather McDermid and Jeff Goldberg, for their time, advice and help. Thanks also to Mike Walter and Joe Culotti for serving as external examiners. I also thank Chris Link and Joel Rothman for their support, advice and encouragement.

I thank my parents, Morris Fidanque Maduro and Marietta Posoncuy Maduro, my sister Melissa Maduro Richer, and the rest of my family. I also thank my wife Gina, for her dedication and companionship, especially during the more difficult and frustrating parts of this work.

Lastly, I wish to thank Giuseppe Verdi, Giacomo Puccini, Gaetano Donizetti and Vincenzo Bellini for writing some of the world's most beautiful operas, and José Carreras, Alfredo Kraus, Jussi Björling and Ben Heppner for singing it.

Table of Contents

1 Introduction

1 Introduction 1

1.1 The nervous system 1
1.2 *C. elegans* as a model system 3
1.3 *C. elegans* development 6
1.4 *C. elegans* behaviours 7
1.5 *C. elegans* musculature 8
1.6 The *C. elegans* nervous system 9

1.6.1 Anatomy 9
1.6.2 Development 11

1.7 Tools for the analysis of the *C. elegans* nervous system 12
1.8 Genetic analysis of the *C. elegans* nervous system 14

1.8.1 The Uncoordinated class of C. elegans mutations 14
1.8.2 unc-86 mutations affect asymmetry in neuroblast lineages 15
1.8.3 unc-4 affects synaptic specificity 16
1.8.4 Synaptic transmission 16
1.8.5 Genes involved in the synthesis and processing of acetylcholine (ACh) 17
1.8.6 Serotonin and egg laying 18
1.8.7 Axonal transport of vesicles by unc-104 19
1.8.8 Synaptic vesicle fusion and release 20
1.8.9 Mutations affecting process placement 21

1.9 Axon guidance 21

1.9.1 Basic forces in axon guidance 21
1.9.2 The Optic Tectum 22
1.9.3 The Ventral Midline 24
1.9.4 The netrin family of guidance molecules and their receptors 24
1.9.5 Genes affecting axon guidance in fascicles 26

1.10 Growth cone motility and axonogenesis 27

1.10.1 The study of axonogenesis 27
1.10.2 Structure of the growth cone 27

1.10.3 Three stages of growth cone motility 28
1.10.4 Genes affecting axonogenesis in *C. elegans* 29

1.11 Background to the thesis 31
1.12 Goals of this thesis 33
Figures 34
1.13 Bibliography 56

2 Identification and Cloning of *unc-119*, a Gene Expressed in the *Caenorhabditis elegans* Nervous System 68

2.1 Introduction 68

2.2 Materials and Methods 70

2.2.1 Strains and Genetics 70
2.2.2 Generation of *unc-119* alleles 70
2.2.3 Mapping of *unc-119* 70
2.2.4 Construction of *unc-119(e2498)*; *daf-7(e1372ts)* and *unc-119(e2498)*; *daf-11(m47ts)* double mutants 72
2.2.5 Phenotypic Analysis 72
2.2.6 Molecular Analysis 73
2.2.7 Microscopy and C. elegans transformation 75

2.3 Results 76

2.3.1 Identification and mutant phenotype of *unc-119* 76
2.3.2 Mapping of *unc-119* 78
2.3.3 Cloning of *unc-119* 79
2.3.4 *unc-119* cDNAs 83
2.3.5 Expression of *unc-119* 86

2.4 Discussion 87
Figures 93
2.5 Bibliography 103
2.6 Addendum 109
2.7 Screen for new *unc-119* mutations 109
2.8 Electron Micrographs of wild type and *unc-119* muscle 110
2.9 Placement of *unc-119* in the dauer formation pathway 111
2.10 Onset of *unc-119*::lacZ expression 111
2.11 Inaccuracies in the *unc-119* sequence 113
2.12 Stem-loop structure in the *unc-119* 5'UTR 113
2.13 Altered *unc-119* mRNA in the *ed9* allele 114
2.14 The molecular identification of the *eDf2* breakpoint 115

2.14.1 Background 115
2.14.2 Cloning of the *eDf2* breakpoint 116
2.14.3 LGIIIR is duplicated CB1517 117
2.14.4 Models for the origin of *eDf2* and *eDp6* 118

2.15 Homologues of UNC-119 119
Tables 120
Figures 121
2.16 Bibliography 132

3 Conservation of function and expression of *unc-119* from two Caenorhabditis species despite divergence of non-coding DNA 134

3.1 Introduction 134

3.2 Materials and Methods 135

3.2.1 Southern Analysis and Library Screening 135
3.2.2 5'RACE Analysis and DNA Sequencing 136
3.2.3 Reporter Gene Construction 137

3.3 Results and Discussion 137

3.3.1 Identification of ~~the C.~~ briggsae homologue of unc-119 137
3.3.2 The coding regions of the homologues are conserved 138
3.3.3 Intron and flanking sequences are divergent 140
3.3.4 Putative promoter elements are conserved in the 5' flanking region 141
3.3.5 Reporter gene expression is conserved between the two species 142
3.3.6 Conclusion 144

Figures 145
3.4 Bibliography 155
3.5 Addendum 158
3.6 Methods 158

3.6.1 Preparation of total RNA 158
3.6.2 Northern hybridization 159
3.6.3 *C. briggsae unc-119*::GFP fusion plasmid 159

3.7 Northern Blot Analysis of *unc-119* 160

3.8 Reporter gene fusion of *C. briggsae unc-119* with GFP 160
3.9 Similarity of C27H5.1 to a vertebrate protein 162
Figures 163
Bibliography 166

4 *C. elegans UNC-119* defines a new class of neural gene involved in axogenesis, and is functionally conserved with a human protein 167

4.1 Introduction 167
4.2 Materials and Methods 170

- 4.2.1 Construction of the *unc-119*::GFP nervous system marker 170
- 4.2.2 Amplification of HRG4 and DmUNC119 cDNAs 171
- 4.2.3 Construction of *unc-119*::HRG4 and *unc-119*::DmUNC119 fusions 172
- 4.2.4 Identification and cloning of *Drosophila melanogaster* homolog 173
- 4.2.5 Construction of transgenic *C. elegans* strains 173
- 4.2.6 Phenotypic assays 174
- 4.2.7 Immunohistochemistry 174
- 4.2.8 *C. elegans* microscopy 175

4.3 Results 175

- 4.3.1 *unc-119* mutants have a presynaptic egg laying defect 175
- 4.3.2 *unc-119* mutants have abnormal chemosensory neurons 177
- 4.3.3 *unc-119* mutants have multiple neuronal structure defects 177
- 4.3.4 A functional UNC-119::GFP fusion is found in axons 180
- 4.3.5 UNC-119 is predicted to have a human homolog based on sequence 181
- 4.3.6 Identification of a novel *Drosophila melanogaster* homolog 181
- 4.3.7 The human homolog HRG4 can provide UNC-119 function in *C. elegans* 182
- 4.3.8 A functional HRG4::GFP fusion is also found in axons 183
- 4.3.9 The carboxy terminus, but not the amino terminus, is required for HRG4 function in *C. elegans* 184
- 4.3.10 The *Drosophila* homolog DmUNC119 can replace

UNC-119 in *C. elegans* 185

4.4 Discussion 185

4.4.1 *unc-119* is required for axon outgrowth and fasciculation 185
4.4.2 UNC-119, HRG4 and DmUNC119 define a new class of proteins 186
4.4.3 Possible UNC-119/HRG4 functions 187

Figures 190
4.5 Bibliography 200
4.6 Addendum 205
4.7 Materials and Methods 205

4.7.1 mRNA analysis in *Drosophila* 205
4.7.2 *C. elegans* UNC-119::GFP fusions 206
4.7.3 *C. briggsae* full-length GFP fusion 207
4.7.4 Construction of pDP#MM065 *unc-119* expression vector 207

4.8 Degenerate primers for identifying UNC-119 homologues 208
4.9 The *Drosophila* DmUNC119 gene 209
4.10 Unusual punctate localization of some *unc-119*::GFP reporters 210
4.11 Structures visualized with *unc-119*::GFP 214
4.12 Visualization of HSN cell bodies in N2 and *unc-119* 215
4.13 Vectors for transgene expression from the promoter 216
4.14 An *unc-119*::GFP fusion with a larger 5' flanking region 217
Figures 218
Bibliography 229

5 Preliminary studies of the UNC-119 protein

5.1 Introduction 230

5.1.1 Previous work 230
5.1.2 Where is the UNC-119 protein found in the cell? 231
5.1.3 When is UNC-119 function required? 232
5.1.4 With what other proteins does UNC-119 interact? 232

5.2 Materials and Methods 233

5.2.1 Plasmid nomenclature 233
5.2.2 Construction of GST fusion plasmids 234
5.2.3 Peptide synthesis and rabbit immunization 235
5.2.4 Spot blot analysis 236
5.2.5 Fixation and antibody staining 236
5.2.6 Construction of HA-tagged *unc-119* 238
5.2.7 Construction of heat-shock *unc-119* fusions 239
5.2.8 *unc-6*::UNC-119 transgene 239
5.2.9 Construction of pUBP1 yeast two-hybrid "bait" plasmid 240
5.2.10 Yeast two-hybrid screen 241

5.3 Results 242

5.3.1 Expression of recombinant UNC-119 and HRG4 in *E. coli* 242
5.3.2 Testing of antisera to two HRG4 epitopes 243
5.3.3 Immunological detection of hemagglutinin (HA) tagged UNC-119 245
5.3.4 Expression of heat-shock driven *unc-119* expression 246
5.3.5 Attempted rescue by an *unc-6*::UNC-119 transgene 248
5.3.6 Screening of a yeast two-hybrid library 249

5.4 Discussion 251

5.4.1 Comments regarding this section 251
5.4.2 Expression of recombinant GST::UNC-119 251
5.4.3 Antibodies to HRG4 synthetic peptides 252
5.4.4 Hemagglutinin tagged UNC-119 253
5.4.5 Heat-shock expression of UNC-119 254
5.4.6 Partial rescue by an *unc-6*::UNC-119 transgene 255
5.4.7 Preliminary yeast two-hybrid screen with UNC-119 255
5.4.8 Concluding remarks 257

Tables 259
Figures 263
5.5 Bibliography 270

6 General Discussion 272

6.1 Introduction 272
6.2 The basis for the *unc-119* mutant phenotype 272

6.2.1 *unc-119* mutations 272
6.2.2 The basis for loss of coordinated movement 273
6.2.3 The dauer formation and amphid dye-filling defects 274
6.2.4 Basis for the egg laying and constitutive pharyngeal pumping defects 274

6.3 Conclusions about *unc-119* gene expression 276

6.3.1 *unc-119* reporter transgenes 276
6.3.2 *unc-119* transcripts 277

6.4 The set of *unc-119* alleles 277
6.5 The partial rescue of *unc-119* by an incomplete *unc-119* gene 280
6.6 Correlation of *unc-119* expression and the mutant phenotype 281
6.7 The functional and structural conservation of UNC-119 283
6.8 A possible role for UNC-119 in growth cone extension 285
6.9 A role for UNC-119/HRG4 in maintenance 289
6.10 Future work on *unc-119* 290
6.11 Concluding remarks 292
Figures 294
Bibliography 296

Appendices 299

Appendix 1: *Drosophila* chromosome *in situ* of DmUNC119 299
Appendix 2: *Drosophila* embryo *in situ* with DmUNC119 300

List of Tables

Table 2.1 Sequencing errors in *unc-119* sequence 120

Table 5.1 Rabbit injections of HRG4 peptides 259

Table 5.2 Summary of staining seen in DP142 with anti-HRG4 antisera 260

Table 5.3 Summary of anti-HA staining in *unc-119*::HA 261

Table 5.4 List of cDNAs obtained in preliminary yeast two-hybrid screen 262

2

List of Figures

Figure 1.1 Appearance of wild type *C. elegans* 34
Figure 1.2 Formation of *C. elegans* founder cells 35
Figure 1.3 Life cycle of *C. elegans* 36
Figure 1.4 Diagram of muscle quadrants of body wall muscle 37
Figure 1.5 The *C. elegans* axon scaffold 38
Figure 1.6 Morphology of *C. elegans* neurons 39
Figure 1.7 Connectivity of egg laying neurons 40
Figure 1.8 Altered Q neuroblast lineage in *unc-86* mutants 41
Figure 1.9 Similar lineages of P neuroblasts 42
Figure 1.10 Locomotory circuit disrupted in *unc-4* mutants 43
Figure 1.11 Synaptic release cycle in *C. elegans* 44
Figure 1.12 The egg laying model 45
Figure 1.13 Muscle arms can find misplaced motor neurons 46
Figure 1.14 Forces that act during axon outgrowth 47
Figure 1.15 Retinotectal projection 48
Figure 1.16 Growth of interneurons in rat spinal cord 49
Figure 1.17 Expression of netrins and their receptors in three systems 50
Figure 1.18 Model for netrin-based growth cone motility 51
Figure 1.19 Schematic depiction of a growth cone 52
Figure 1.20 Model for growth cone motility 53
Figure 1.21 Site selection and site stabilization 54
Figure 1.22 Genetic and physical maps around *unc-119* 55

Figure 2.1 Wild type, *unc-119* and *eDf2*; *eDp6* animals 93

Figure 2.2 Quantitation of behavioural defects 94

Figure 2.3 Genetic map around *unc-119* 95

Figure 2.4 Southern analysis places cosmids at *eDf2* breakpoint 96

Figure 2.5 Physical map of the *unc-119* region 97

Figure 2.6 Southern blots show an *e2498*-specific polymorphism 98

Figure 2.7 Short-range physical map and cDNA of *unc-119* 99

Figure 2.8 Sequence of the *unc-119* gene 100

Figure 2.9 X-gal staining of *unc-119*::lacZ animals 102

Figure 2.10 Mutagenesis scheme for generating new *unc-119* alleles 121

Figure 2.11 Electron micrograph of N2 muscle sarcomere 122

Figure 2.12 Electron micrograph of *e2498* muscle sarcomere 123

Figure 2.13 Placement of *unc-119* in dauer formation epistatic pathway 124

Figure 2.14 Diagram of nuclei in an early *unc-119*::lacZ embryo 125

Figure 2.15 Stem-loop structure in *unc-119* 5'UTR 126

Figure 2.16 Sequence alterations in *unc-119* mutants 127

Figure 2.17 Localization of the *eDf2* breakpoint 128

Figure 2.18 Sequences at the *eDf2* breakpoint 129

Figure 2.19 Southern hybridization of *eDf2* and *eDp6* 130

Figure 2.20 Models for the origin of *eDf2* and *eDp6* 131

Figure 3.1 Southern analysis of a *C. briggsae* λ clone 145

Figure 3.2 Restriction map, rescuing clones and gene arrangement of *C. briggsae unc-119* gene 146

Figure 3.3 Quantitation of rescue of *C. elegans unc-119* by the *C. briggsae* gene 147

Figure 3.4 Sequence of the *C. briggsae unc-119* gene 148

Figure 3.5 Alignment of *C. elegans* and *C. briggsae* UNC-119 150

Figure 3.6 Matrix comparison of *C. elegans* and *C. briggsae unc119* genes 151

Figure 3.7 X-gal staining of *C. briggsae unc-119*::lacZ 154

Figure 3.8 Run-off templates for synthesis of antisense RNA 163

Figure 3.9 Northern hybridization of *unc-119* in *C. elegans* and *C. briggsae* 164

Figure 3.10 Fluorescence microscopy of *C. briggsae unc-119*::GFP expression 165

Figure 4.1 Quantification of *unc-119* defects and rescue by an *unc-119*::HRG4 transgene 190

Figure 4.2 Video images of *unc-119* strains 191

Figure 4.3 Uptake of FITC into sensory neurons 192

Figure 4.4 *unc-119*::GFP fluorescence in various Unc mutants 193

Figure 4.5 Abnormal processes in *unc-119* mutants 194

Figure 4.6 Detailed views of *unc-119* axon defects 196

Figure 4.7 Localization of functional UNC-119::GFP and HRG4::GFP 197

Figure 4.8 Comparison of UNC-119, HRG4 and DmUNC119 198

Figure 4.9 Rescuing ability of various *unc-119*::HRG4 transgenes 199

Figure 4.10 Degenerate oligonucleotides for PCR 218

Figure 4.11 Arrangements of *C. elegans*, *C. briggsae* and *Drosophila* genes 220

Figure 4.12 Partial sequence of *Drosophila* DmUNC119 gene 221

Figure 4.13 Punctate localization of functional UNC-119::GFP 223

Figure 4.14 Comparison of various *unc-119*::reporter fusions 224

Figure 4.15 Specific neurons expressing *unc-119*::GFP 225

Figure 4.16 Accumulation of *unc-119*::GFP in particular cells 226

Figure 4.17 Staining of HSNs in wild type and *e2498* 227

Figure 4.18 *unc-119* expression vectors 228

Figure 5.1 Diagram of GST fusion plasmids 263

Figure 5.2 Diagram of heat-shock fusion plasmids 264

Figure 5.3 Features of GAL4::UNC-119 "bait" plasmid 265

Figure 5.4 Amino acid sequences of NIHPW6 and NIHPW7 266

Figure 5.5 Fluorescence microscopy of preliminary antibody staining with rabbit serum 267

Figure 5.6 Fluorescence of heat-shock induced UNC-119::GFP 269

Figure 6.1 Proposed basis for uncoordinated locomotion of *unc-119* mutants 294

Figure 6.2 Possible model for UNC-119 function 295

1 Introduction

1.1 The nervous system

The nervous system of an organism gives it the capacity to sense its environment, both external and internal, and carry out actions that will favour its survival. Sensory organs obtain information for various modalities, such as sight, sound, taste, and temperature; neural centres (ganglia) coordinate and interpret this information, and send appropriate signals to mediate behavioural or physiological responses. In higher animals, the brain permits conscious thought, emotions and the retention of complex memories and experiences. While functionally complex, the extraordinary information processing abilities of the brain arise from intricate connections among some 10^{10}-10^{12} individual neurons, that in turn make use of only a small number of stereotyped signals (Nicholls *et al.*, 1992). A central goal of biology is to understand the development and function of a complete nervous system, with a long-term view towards comprehending the nature of the human brain. An understanding of the fundamental principles underlying neuron development and function is a prerequisite for the study of higher brain functions such as perception, learning, behaviour and consciousness.

An important finding in this century has been that the basic mechanisms of neuron function are remarkably similar, both among the different neuron types within an individual, and from species to species (Nicholls *et al.*, 1992). This has permitted neurobiologists to draw general conclusions about basic physiological phenomena, such as synaptic transmission and the propagation of electrical signals, and more complex phenomena, such as coordination of locomotion and potentiation of memory, through the study of model organisms. For example, Hodgkin and Huxley (1952) made landmark descriptions of membrane conductance using the giant axon of the squid *Loligo*; the simple nervous system of the leech has allowed descriptions of neural connections involved in its swimming behaviour (Kristan, 1983); and the gastropod *Aplysia* has served as a useful model for the study of habituation (Kandel, 1976). Invertebrates amenable to genetic analysis are also included among model systems used for behavioural studies. Short-term memory has been demonstrated in the fruit fly *Drosophila* (Tully *et al.*, 1994), and the nematode *C. elegans* exhibits classical habituation (Wicks and Rankin, 1996).

Mechanisms of neural development also share common features. The fate of most neurons and glial cells in vertebrates and complex invertebrates is largely determined by inductive interactions (Nicholls *et al.*, 1992). However, this is not true for invertebrates like *C. elegans*, where neural cell fates are determined largely by lineage (Chalfie and White, 1988). Despite this difference, *C. elegans* has emerged as a powerful system for studying a complete nervous system, through the exploitation of its simplicity coupled with the power of genetic analysis (Brenner, 1974). Molecular characterization of *Caenorhabditis* neural mutants has led to the discovery of proteins with vertebrate homologues, confirming the evolutionary conservation of neural development and function (e.g. Tessier-Lavigne and Goodman, 1995; Südhof, 1996).

This thesis describes the identification of a novel factor, UNC-119, that participates in the neural development of *C. elegans*. The *unc-119* gene has functionally conserved homologues in other metazoans, including humans, consistent with a conserved biological function. This study, the first to examine *unc-119*, focuses primarily on a molecular genetic approach towards understanding its function, expression and conservation. The first parts of the thesis describe the molecular identification of the gene and its expression in

Caenorhabditis. The mutant phenotype is then described at a cellular level using a novel nervous system marker. The sequence of an apparent human homologue, HRG4, is used to identify a homologue from the fruit fly *Drosophila*. Both of these gene products are shown to function in place of UNC-119 in *C. elegans*. Finally, attempts to detect UNC-119 *in situ* are described. In the general discussion, models for possible biochemical roles of UNC-119 are developed. Although a concrete role for UNC-119 is not certain, a function in axon outgrowth is proposed.

1.2 *C. elegans* as a model system

Since first introduced as a genetic system more than a quarter-century ago (Brenner, 1974), the nematode *C. elegans* has emerged as one of the premiere models for studying the development and function of a complete nervous system. For this analysis, *C. elegans* has a number of features that make it advantageous. The animals are small, ~1 mm in length when fully grown, and have a short generation time, three days at 25°C. They are transparent at all developmental stages, facilitating whole animal analysis by differential interference contrast microscopy (Nomarski, 1955) or fluorescence microscopy, and are easily propagated on small agar plates, fed a diet of *E. coli*. Viable stocks can be recovered after storage in liquid N_2 or at -80°C.

Animals exist in two sexes, males and hermaphrodites (Figure 1.1). The hermaphrodite, chromosomally XX, can be considered a female that manufactures a small amount of sperm early in sexual maturity, that it can use throughout the adult lifespan for self-fertilization. Males, possessing a single X chromosome (XO), can arise from spontaneous X-chromosome nondisjunction and can be used to cross-fertilize hermaphrodites. Mutations are readily made using common mutagens such as ethyl methane sulfonate (EMS), that produces forward mutations in an average-sized gene at a rate of ~10^{-4} per generation (Brenner, 1974). There are five pairs of autosomes in addition to the sex chromosomes, and many morphogenetic markers, useful for mapping, have been described (Brenner, 1974; Wood, 1988; Hodgkin, 1997).

Another aspect of *C. elegans* biology that has made it useful for cytological study is its nearly invariant cell lineage, coupled with a small number of nuclei (approximately 1000) in the adult: each zygote follows a stereotyped spatial and temporal pattern of cell divisions that result in a constancy of cell number (eutely), and constancy of cell position, from animal to animal (Sulston and Horvitz, 1977; Sulston *et al.*, 1983). Knowledge of the cell lineage and the ability to identify individual cells have been important for understanding gene spatial and temporal expression patterns, and for dissecting developmental mechanisms. Experimental alterations in the lineage through mutation, embryo manipulation, or by laser ablation of specific cells, have allowed a detailed characterization of many developmental processes, such as establishment of embryonic polarity (e.g. *par* genes; Kemphues *et al.*, 1988, Kirby *et al.*, 1990), induction of cell fate by cell-cell interaction (e.g. *glp-1*; Moskowitz *et al.*, 1994), asymmetrical cell division (e.g. *vab-3*; Chamberlin and Sternberg, 1995), cell death (Driscoll, 1992), genetic control of aging (Kenyon, 1997) and neural pathfinding (see below).

In the mid 1980's, an ambitious project was initiated to determine the entire DNA sequence of all 100 megabases of the *C. elegans* haploid genome (Sulston *et al.*, 1992). The initial stages of the project involved the assembling of large collections of yeast artificial chromosome (YAC) and cosmid clones into contiguous regions to establish a physical map (Coulson *et al.*, 1986, 1988, 1991), an important resource for positional cloning (and that was used for the cloning of *unc-119*; see Chapter 2). Animals can be easily made transgenic by microinjection of cloned DNA into the syncytial gonad, an approach used to verify the cloning of a gene by transformation rescue (Fire, 1986; Mello, 1992). Genes predicted from the *C. elegans* genomic sequence (using GENEFINDER; described in Waterston *et al.*, 1997) can be confirmed by correlation with an expressed-sequence tag (EST), conservation with known genes using database search algorithms such as BLAST (Altschul *et al.*, 1997), and by comparison with homologous sequences from the related nematode *C. briggsae* (Waterston *et al.*, 1997). The anticipated completion date of the *C. elegans* genome sequencing project is late 1998. The physical and genetic maps, as well as partial and completed DNA sequences, EST data, and expression patterns of some genes are available to the scientific community through the Internet (Waterston *et al.*, 1997).

The ability to predict gene structure directly from genomic DNA sequence has led to the development of approaches for reverse genetics. A 'sib-selection' screen for spontaneous transposon Tc1 insertions in a predicted gene can be performed on pooled nematode cultures by the polymerase chain reaction (Plasterk, 1995). The Tc1 insertions can be used as a starting strain for the generation of imprecise excisions to create deletion alleles, or for gene replacement from a homologous transgene supplied *in trans* (Plasterk and Groenen, 1992). More recently, this method has been superseded by a direct EMS or trimethylpsoralen (TMP) induced mutagenesis approach, that can produce deletion alleles in a single screen (Yandell *et al.*, 1994; Jansen *et al.*, 1997). An alternative to a chromosomal disruption is to phenocopy loss of function by gonadal microinjection of either sense or anti-sense RNA corresponding to the exonic regions of a predicted gene: a proportion of the progeny may show the phenotype (e.g. Guo and Kemphues, 1995). Many maternal genes and some zygotic genes are amenable to this approach, but the mechanism of RNA-induced phenocopy, and its apparent heritability, is unknown (C. Mello, pers. comm.).

1.3 *C. elegans* development

The developmental lineage of cells in *C. elegans* is nearly invariant (Wood, 1988). Most cell divisions occur during embryogenesis, which is comprised of two phases: (a) cell proliferation and organogenesis, and (b) morphogenesis. The zygote undergoes several asymmetric divisions to give rise to the six founder cells AB, MS, E, C, D and P_4 (Figure 1.2). Gastrulation occurs with the inward entry of the anterior and posterior daughter cells of E moving from the ventral surface of the embryo after approximately 100 minutes. During the remainder of the proliferative phase, cells follow a stereotyped pattern of divisions, migrations and cell deaths. Approximately seven hours post-fertilization, elongation, connection of neurons, and cuticle secretion occurs. The first-stage larva, the L1, hatches approximately 14 hours after fertilization, with 958 cells in the hermaphrodite and 960 cells in the male (Wood, 1988).

The animal progresses through three additional larval stages (L2, L3 and L4) before forming the adult (Figure 1.3). Each transition is accompanied by a molt: a new cuticle is generated below the old cuticle, and the old cuticle is subsequently shed. Additional post-embryonic developmental changes do occur, which include modifications to the nervous system, the generation of the gonad, and the formation of sexually dimorphic tissues: the hermaphrodite vulva, and the specialized male tail (Sulston and Horvitz, 1977).

At a temperature of 20°C, hermaphrodites and males will live for approximately 17 days (Wood, 1988), although different strains can vary from 12 to 18 days (Kenyon, 1997). In an unfavourable growth environment, larvae can enter an alternative L3 stage, the "dauer" (indicated in Figure 1.3). Formation of dauer larvae is favoured under conditions of low food, high temperature and overcrowding, mediated by chemosensation of the environment. Animals stop feeding, accumulate storage granules, and secrete a specialized cuticle that is resistant to dessication; in this state, they can persist for several months (Kenyon, 1997). Under favourable conditions, animals exit the dauer state and resume growth, with a similar life span to animals that have progressed through a normal L3. Mutations that affect dauer formation result in either inappropriate entry into dauer (Daf-c mutations) or resistance to dauer formation (Daf-d); most of the Daf genes have been placed in a genetic pathway (summarized in Riddle and Albert, 1997). It is possible that dauer larvae may activate a longevity program, as mutations in the dauer formation genes *daf-2*, *age-1* and *daf-28* can increase adult lifespan (Malone *et al.*, 1996; Kenyon, 1997).

1.4 *C. elegans* behaviours

Despite a small number of cells and a simple nervous system (described in the next section), *C. elegans* displays a diversity of behaviours. Animals propel themselves backwards or forwards through the propagation of dorsal-ventral sinusoidal waves along the body. They are capable of sensing and reacting to many environmental stimuli, such as temperature, osmolarity, and chemical odorants (Wood, 1988). If touched at the head or tail, animals will recoil and move in the opposite direction (tap withdrawal), or may even retract from a precipice (Wood, 1988). There is evidence that *C. elegans* neurons may not be exclusively involved in a particular behaviour: mutations affecting chemosensation have shown that the ability to sense a variety of stimuli may result from overlapping sensory neuron function (Bargmann and Horvitz, 1991a).

Other behaviours that have been studied include pharyngeal pumping (Avery, 1993), defecation (Thomas, 1990), hermaphrodite egg laying (Desai *et al.*, 1988; Desai and Horvitz, 1989; Weinshenker *et al.*, 1995), and male mating (Liu and Sternberg, 1995). Animals also exhibit classical habituation to the tap-withdrawal response (Wicks and Rankin, 1996).

1.5 *C. elegans* musculature

The most prominent and best characterized set of muscles in *C. elegans* are those of the body wall. The body wall muscles are arranged in four strips arranged longitudinally, each occupying a quadrant of the animal as seen in cross section (Figure 1.4). Each quadrant consists of 24 cells along the body, except for the ventral left quadrant, which contains 23 (Sulston and Horvitz, 1977). Each cell consists of a cell body, containing the nucleus and cytoplasmic organelles, the spindle, which houses the myofilament lattice, and a muscle arm, a process extending to the dorsal or ventral nerve cord, which allows muscle cells to receive synaptic input from motor neurons. This arrangement is unusual among metazoans: in most animals, motor neurons leave nerve fascicles to innervate muscle (Waterston, 1988).

The myofilament lattice itself is similar to vertebrate muscle, consisting of myosin-containing thick filaments and actin-containing thin filaments; however, unlike vertebrates, the muscle lattice is obliquely striated, and is laterally connected directly to the cuticle along its length (Waterston, 1988). An electron micrograph of a body wall muscle sarcomere in transverse section is presented in Chapter 2 (Figure 2.11).

The arrangement of the body wall muscles, and their pattern of innervation as dorsal or ventral left-right pairs, allows the animal to propagate dorsal-ventral sinusoidal waves along the body for locomotion. Mutations affecting body wall musculature therefore produce paralyzed or uncoordinated animals (see §1.8 below). For example, the gene *unc-54* encodes the major myosin isoform (myoB) of *C. elegans*, and was the first myosin gene to be cloned from any organism (MacLeod *et al.*, 1981). Homozygosity for *unc-54* null mutations results in near total paralysis in the adult, caused by an abnormal myofilament lattice with only ~1/4 the number of thick filaments found in wild type (Waterston, 1988).

In addition to those of the body wall, there are muscles associated with the digestive tract (pharynx, intestine and anus), and those that are sexually dimorphic, the hermaphrodite egg laying muscles and male tail muscles (Waterston, 1988).

1.6 The *C. elegans* nervous system

1.6.1 Anatomy

The nearly invariant cell lineage, and the constancy of cell number and placement, have allowed the precise identification of all cells of the *C. elegans* nervous system (White *et al.*, 1986). The adult hermaphrodite consists of 959 somatic nuclei. Of these, there are 302 neurons and 56 glial and associated support cells. The majority of neural processes form on the basal surface of the body wall, while a small number (20) forms on the basal surface of the pharynx. As such, the *C. elegans* nervous system is comprised of two distinct units (White *et al.*, 1986; Wadsworth and Hedgecock, 1996).

The nervous system is organized into cell body clusters (ganglia) and process bundles, which run longitudinally or circumferentially (Figure 1.5). At the anterior end of the animal is the major neuropil ganglion structure, the nerve ring, which contains a large number of neural cell bodies and interneurons (approximately 100; White *et al.*, 1986). Most neurons have a simple monopolar morphology, possessing few branches, if any (Figure 1.6; White *et al.*, 1986). Aside from the nerve ring, the loose clustering of cell bodies into ganglia does not seem to correlate with functional similarity among the neurons within them (Chalfie and White, 1988). The small number of neurons has allowed the serial electron micrograph reconstruction of the entire nervous system, and a description of the ~7000 chemical synapses and gap junctions made (White *et al.*, 1986). An example is shown in Figure 1.7, which describes the connectivity of neurons in the egg-laying circuitry.

The most prominent longitudinal process bundle is the ventral nerve cord (VNC), which runs along the midline of the animal along the ventral hypodermal ridge (Figs. 1.5, 1.6A; White *et al.*, 1976, 1986; McIntire *et al.*, 1992). The VNC consists of two fascicles that run in parallel: the smaller left fascicle consists of four processes, while the right fascicle contains 40 axons (White *et al.*, 1986; Wightman *et al.*, 1997). The VNC contains axons from neurons whose cell bodies are located in the nerve ring and posterior ganglia, from some laterally placed cell bodies, and from motor neurons with cell bodies adjacent to the right process bundle (White *et al.*, 1986; McIntire *et al.*, 1992).

Despite the simplistic organization of its nervous system, *C. elegans* displays a surprising diversity in the properties of individual neurons. Based on their position, the 302 neurons can be assigned to 118 classes (White *et al.*, 1986). Neurotransmitters found in many other organisms have been demonstrated in *C. elegans*, such as acetylcholine, GABA (γ-aminobutyric acid), dopamine, serotonin, and glutamate; about half of the *C. elegans* neurons have been assigned neurotransmitters (summarized in Rand and Nonet, 1997a).

1.6.2 Development

Most of the nervous system is generated from descendants of the AB blastomere; the remainder arise from C and MS (Figure 1.2; Chalfie and White, 1988). The majority of neurons (approximately 220) are generated during the proliferative first phase of embryogenesis (Sulston, 1976). An additional period of neurogenesis occurs post-embryonically, when the majority of ventral nerve cord motor neurons are generated. There are eight classes of motor neuron (DAn, DBn, DDn, VAn, VBn, VCn, ASn, VDn) in the adult; only the DAn, DBn and DDn arise embryonically. The development of the remaining classes occurs towards the end of the L1 stage (Sulston and Horvitz, 1977; Sulston *et al.*, 1983), and is accompanied by a rewiring of the DD motor neurons: At hatching, the DD cells receive synapses from dorsal cord DA/DB neurons, and synapse onto muscles in the ventral nerve cord; in the adult, they have neuromuscular junctions in the dorsal cord, and receive input from VA/VB neurons in the ventral nerve cord (White *et al.*, 1978).

In addition to the lineage, establishment of the nervous system also relies on some specific cell deaths, as occurs in the neuroblast lineages that generate the late-arising motor neurons (Chalfie and White, 1988), and on cell migrations, as with the hermaphrodite-specific neurons (HSNs; Garriga and Stern, 1994). Lastly, the development of mating-specific neurons accompanies sexual differentiation in males during the L3 stage, resulting in an extended nervous system: while adult hermaphrodites have 302 neurons and 56 glial and associated support cells, males have 381 neurons and 92 glial and supporting cells (Chalfie and White, 1988).

Differentiation of neurons requires the precise extension of axons to establish neural connectivity. Mechanisms governing axon guidance are described in §1.9 below.

1.7 Tools for the analysis of the *C. elegans* nervous system

Developmental studies of *C. elegans*, which include elucidation of the cell lineage and development of techniques to analyze animals at the cellular level, have created a set of tools for analysis of the *C. elegans* nervous system. These have collectively emerged as one of the most powerful set of methods available for the study of any nervous system.

As mentioned earlier, the transparency of animals, the knowledge of the cell lineage and the establishment of neural connectivities allows unique identification of neurons, and, therefore, a precise description of mutant phenotypes affecting the nervous system. Ablation of specific cells with a laser microbeam can identify neurons required for a particular behaviour, such as pharyngeal pumping (Avery and Horvitz, 1989), or for a developmental decision, such as entry into the dauer larva resting state (Perkins *et al.*, 1986; Bargmann and Horvitz, 1991b). The size of *C. elegans* animals also permits ultrastructural examination of neuron structure by electron microscopy (e.g. Perkins *et al.*, 1986).

The projection of sensory cilia into the worm external microenvironment allows the uptake of fluorescent dyes such as DiO and FITC into chemosensory neurons (Hedgecock *et al.*, 1985; Perkins *et al.*, 1986; Starich *et al.*, 1995). These enable the cell bodies of the amphid and phasmid neurons to become fluorescent, by diffusion of the dye through the cilia and along the axons; this ability depends upon the correct structure of the sensory cilia (Perkins *et al.*, 1986).

Animals can respond to stimulation by various pharmacological agents, such as levamisole, an acetylcholine agonist (Lewis *et al.*, 1980), or neurotransmitters such as serotonin (Desai *et al.*, 1988). The restriction of particular neurotransmitters to a subset of neurons in *C. elegans* has allowed the examination of neuron structure in fixed animals by immunohistochemical analysis. For example, a subset of commissures can be revealed by detection of the neurotransmitter GABA (McIntire *et al.*, 1993), facilitating characterization of mutations affecting axonal outgrowth (McIntire *et al.*, 1992). Similarly, the ease of making transgenic animals has allowed the preliminary characterization of neural gene expression using reporter gene fusions to *lacZ*, and can provide a method by which to detect specific neurons (e.g. *ceh-23::lacZ*; Wightman *et al.*, 1996). Alternatively, neural gene products may be expressed ectopically by placing a coding region under control of a different promoter (e.g. Hamelin *et al.*, 1993).

Recently, the availability of the green fluorescent protein (GFP) from the jellyfish *Aequoria victoria*, the subsequent discovery of a mutation that enhances its fluorescence, and the availability of a set of useful fusion vectors, have allowed visualization of neural processes in living animals using cloned neural-specific promoters (Chalfie *et al.*, 1994; Heim *et al.* 1995; A. Fire, J. Ahnn, G. Seydoux and S. Xu, pers. comm.). The main advantage of GFP reporter fusions is that neurons can be followed during development, at much higher resolution, without the need for destructive fixation or detection methods. Furthermore, inclusion of appropriate signal sequences, or production of fusion gene products, can allow subcellular localization of GFP fluorescence. For example, this has been done for synaptobrevin, a component of synaptic vesicles (Rand and Nonet, 1997b). The coupling of GFP fusions and newer fluorescence imaging techniques, such as computer-controlled confocal laser scanning microscopy, will eventually permit four-dimensional (x, y, z axes and time-lapse) analysis of neurons to be made.

Lastly, methods have been developed that allow electrophysiological analysis of neural function and synaptic transmission (Avery *et al.*, 1995; Lockery and Hall, 1995). For example, the pharynx can be introduced into a suction pipet to allow measurement of current flow associated with contraction of the pharyngeal muscles (Avery *et al.*, 1995). The distantly-related nematode *Ascaris* shares a number of nervous system structural features with *C. elegans*, such as the same number and organization of ventral nerve cord motor neurons (Stretton *et al.*, 1978). As such, studies in *C. elegans* can be complemented by direct physiological analysis of homologous *Ascaris* neurons (Chalfie and White, 1988): due to its larger size, *Ascaris* is more amenable to this type of manipulation.

1.8 Genetic analysis of the *C. elegans* nervous system

1.8.1 The Uncoordinated class of C. elegans mutations

The physical and genetic independence of pharyngeal neurons from the remainder of the nervous system allows *C. elegans* to continue feeding, even if defects in the remainder of the nervous system cause body paralysis. Only the CANL/CANR neuron pair, in the body, and M4 in the pharynx, are required for development (Chalfie and White, 1988). As hermaphrodite fertilization is internal, mutant stocks can be propagated in the absence of mating. This allows the recovery of homozygous stocks where animals show severe neural defects (as well as other movement-compromised animals, such as those with abnormal body wall musculature), facilitating the isolation of a wide variety of nervous system mutations.

Most animals that have varying degrees of paralysis or irregular body movement are grouped under the Uncoordinated class of mutants (Brenner, 1974), of which ~130 are currently known (Hodgkin, 1997). Others with Unc phenotypes have been named through homology with other known factors (e.g.

snt-1, the *C. elegans* homologue of synaptotagmin; Nonet *et al.*, 1993) or for other aspects of the phenotype, such as a lineage defect (e.g. *lin-6*; Sulston and Horvitz, 1981) or a specific behavioural/pharmacological defect (e.g. *ric-3*; resistance to inhibitors of cholinesterase; Nguyen *et al.*, 1995). Approximately 65% of Unc genes are known to have neural defects, while most of the remainder have muscle defects (Hodgkin, 1997).

Genes that mutate to an Unc phenotype as the result of a neural defect can be grouped into four general classes: mutations affecting the cell lineages producing motor neurons, those that alter synaptic specificity, those that affect neurotransmitter function, and mutations affecting process placement (Chalfie and White, 1988). What follows below are examples of mutations affecting each class, with particular emphasis on synaptic transmission and axonal outgrowth.

1.8.2 unc-86 mutations affect asymmetry in neuroblast lineages

Animals mutant for *unc-86* have neural defects resulting from abnormalities in many cell lineages producing neurons, including those that are interneurons, sensory neurons and motor neurons (Chalfie *et al.*, 1981). The *unc-86* gene product, a transcription factor of the POU-domain class, is required for particular daughter cells to become distinct from their mother cells, and for some neurons to differentiate (Desai *et al.*, 1988; Finney *et al.*, 1988; Finney and Ruvkun, 1990). In the lineages affected in *unc-86* mutants, one daughter cell reiterates the lineage normally associated with its mother cell (Chalfie *et al.*, 1981). In wild types, expression of UNC-86 occurs in the daughter cells affected in *unc-86* mutants (Finney and Ruvkun, 1990; Ruvkun, 1997). This suggests that the UNC-86 protein is an early global regulator in related neuroblast lineages, acting downstream of asymmetrically segregated factors. Figure 1.8 shows an example of a reiterated lineage in an *unc-86* mutant for the neuroblast Q.

1.8.3 unc-4 affects synaptic specificity

The late-arising motor neurons descend from a set of 13 cells, the P neuroblasts, which follow a characteristic pattern of division and differentiation (with some modifications), each generating a similar set of distinctly different neurons in parallel, as summarized in Figure 1.9 (Chalfie and White, 1988). The use of reiterated lineage patterns suggests that generation of neuron types within these sublineages relies on cell-autonomous signals.

For example, the *unc-4* gene is required for the correct wiring of three of the VA neurons generated from lineages where the sister of VA is VB (White *et al.*, 1992; Miller *et al.*, 1992). In *unc-4* mutants, the VA2, VA3 and VA10 neurons acquire the VB-type presynaptic specificity of their sister cells, resulting in defective backward movement (Figure 1.10; White *et al.*, 1992). The *unc-4* gene product, a Prd/Pax-like homeodomain protein, is expressed in the VA neurons themselves, and as a transcription factor, likely regulates the expression of other genes involved in establishing synaptic specificity in VA2, VA3 and VA10 (Miller *et al.*, 1992; Miller and Niemeyer, 1995).

1.8.4 Synaptic transmission

Communication between neurons or between neurons and an effector cell occurs by synaptic transmission. Of the over 7000 synapses in *C. elegans*, approximately 2000 are neuromuscular junctions, 5000 are chemical synapses between neurons, and 700 are gap junctions (White, 1986; Rand and Nonet, 1997b). Genetic analysis in *Drosophila* and *C. elegans* has identified components involved in synaptic transmission that have escaped biochemical or physiological detection in other systems. In *C. elegans*, in particular, the ability to propagate most stocks with severe neurological defects has certainly been an enormous advantage.

Neurotransmitters are manufactured by a presynaptic cell and packaged into specialized compartments, synaptic vesicles. These vesicles fuse with the plasma membrane, releasing their contents into the synaptic cleft. Specialized receptors on the receiving cell bind the transmitter and initiate a post-synaptic response; this may involve an influx of ions, or the activation of an intracellular signal transduction cascade. A diagram showing the synaptic release cycle is shown in Figure 1.11.

Some mutations affecting the synaptic vesicle cycle can be obtained by selecting for resistance to inhibitors of acetylcholinesterase, such as aldicarb (a Ric phenotype, for resistant to inhibitors of AChE). Wild types exposed to AChE inhibitors become hypercontracted due to toxic overaccumulation of ACh; mutations reducing the buildup of ACh in the synaptic cleft can alleviate the resultant toxicity (summarized in Rand and Nonet, 1997).

1.8.5 Genes involved in the synthesis and processing of acetylcholine (ACh)

The primary excitatory neurotransmitter in *C. elegans* is acetylcholine, based on pharmacological, biochemical and genetic evidence (Rand and Nonet, 1997). ACh is synthesized in the cytoplasm from choline and acetyl coenzyme A, by the enzyme choline acetyltransferase (ChAT); it is degraded in the synaptic cleft by acetylcholinesterase (AChE; Nicholls *et al.*, 1992). Viable mutants for the *C. elegans* ChAT gene *cha-1* are uncoordinated, Ric, small, and slow growing, while homozygotes for null alleles arrest as L1 larvae (Rand, 1989). Most of the cells expressing *cha-1* are motor neurons, consistent with an Unc phenotype (J. Duerr, cited in Rand and Nonet, 1997).

Loading of acetylcholine into vesicles is accomplished by a vesicle acetylcholine transporter (VAChT), the product of the *unc-17* gene (Alfonso *et al.*, 1993). The *cha-1* and *unc-17* loci share the same promoter and first exon; the remainder of the *unc-17* transcript is contained with the first large exon of the *cha-1* gene (Alfonso *et al.*, 1994). A similar gene arrangement occurs in vertebrates (Erickson *et al.*, 1994), suggesting that this structure may be important for regulation, and that this cholinergic locus evolved from a common ancestral form (Rand and Nonet, 1997).

C. elegans has at least three acetylcholinesterase genes with overlapping function, named *ace-1* through *ace-3* (reviewed in Rand and Nonet, 1997). Single and double mutants have few to mild defects, while *ace-1*; *ace-2*; *ace-3* triple mutants arrest after hatching (Johnson *et al.*, 1988).

The postsynaptic response to acetylcholine in *C. elegans* appears to be mediated, at least in part, by ligand-gated receptors of the nicotinic acetylcholine receptor (nAChR) family (Rand and Nonet, 1997). The drug levamisole is a potent nAChR agonist; wild-type animals hypercontract in the presence of the drug, and eventually die (Lewis *et al.*, 1980). This has allowed the identification of genes involved in cholinergic transmission, through selection of mutants resistant to levamisole. Mutations in the genes *unc-29*, *unc-38* and *lev-1*, which encode three nAChR subunits, confer resistance to the drug (Lewis *et al.*, 1987).

1.8.6 Serotonin and egg laying

The egg-laying behaviour of *C. elegans* hermaphrodites has been studied in detail (reviewed in Chalfie and White, 1988). Egg laying is mediated by the vulval and uterine muscles, innervated by the two HSN neurons and six VC neurons; the connectivity of these neurons is known (see Figure 1.7; White *et al.*, 1986). Ablation experiments have shown that the HSNs are required for egg laying, and that conversely, animals lacking the VC neurons can still lay eggs. The HSNs stain with anti-serotonin antibodies (Desai, 1988), and addition of exogenous serotonin, even in the absence of functioning HSNs, results in potentiation of egg laying (Trent *et al.*, 1983). Imipramine can also stimulate egg laying, likely by inhibition of serotonin reuptake (Chalfie and White, 1988).

Mutations have been isolated that result in defective egg laying (an Egl-d phenotype). This phenotype can be classified into four basic types, depending on whether worms are resistant or sensitive to serotonin or imipramine (Trent *et al.*, 1983). For example, mutants for *egl-1*, which results in the inappropriate programmed cell death of the HSNs, lay eggs in response to exogenous serotonin, but not imipramine (Trent *et al.*, 1983). This is consistent with a model in which the HSNs release serotonin, which acts on the egg laying muscles. It has been subsequently shown that acetylcholine is required for egg laying, and that addition of fluoxetine (a selective serotonin reuptake inhibitor, or SSRI), also potentiates endogenous serotonin (Weinshenker *et al.*, 1995). The model for egg laying is shown in Figure 1.12.

1.8.7 Axonal transport of vesicles by unc-104

Transport of vesicles from the cell body to the axon terminal is accomplished by motor proteins of the kinesin family, which anchor to vesicles and move them along microtubule tracks (Rand and Nonet, 1997). Mutations in the gene *unc-104* result in animals with defective motor behaviours, including hatching, locomotion, feeding and defecation (Hall and Hedgecock, 1991). In severe alleles, most or all classes of chemical synapses, including neuromuscular junctions, are rare or absent. There are few synaptic vesicles in axons; instead, vesicles accumulate in neural cell bodies (Hall and Hedgecock, 1991). The *unc-104* gene product is a 180-kD kinesin-related protein (Otsuka *et al.*, 1991), suggesting that UNC-104 is required for synaptic vesicle transport, and that the behavioural phenotypes result from defective synaptic transmission (Hall and Hedgecock, 1991).

Muscle arms are also abnormal in *unc-104* animals. For example, dorsal muscles, which normally make neuromuscular junctions in the dorsal nerve cord, instead extend muscle arms along the lateral hypodermis to the vesicle-rich cell bodies (Figure 1.13; Hall and Hedgecock, 1991). Muscles can therefore adapt to changes in motor neuron structure, suggesting that muscle arm guidance may be mediated by response to chemotropic molecules associated with vesicles (Jorgensen and Rankin, 1997).

1.8.8 Synaptic vesicle fusion and release

Following translocation of synaptic vesicles to the axon terminal and loading of neurotransmitter, synaptic vesicles fuse with the presynaptic membrane in three steps: docking, priming and fusion (Figure 1.11; reviewed in Südhof, 1995).

Null mutations in *unc-18* cause severe paralysis and exhibit resistance to the acetylcholinesterse inhibitor aldicarb (Hosono *et al*., 1987). The product of the *unc-18* gene is proposed to regulate the formation of a stable vesicle-membrane complex (Südhof, 1995; Rand and Nonet, 1997). Consistent with a conserved role, a vertebrate UNC-18 homologue can interact with membrane proteins during vesicle docking (e.g. Hata *et al.*, 1993).

Vesicle fusion is thought to result from the rapid influx of calcium (Südhof, 1995). Synaptotagmins are thought to function as sensors in this step, partially because of evidence that synaptotagmins bind Ca^{2+} cooperatively and undergo a conformational change (Südhof, 1995). The synaptotagmin homologue has been identified in *C. elegans* (*snt-1*; Nonet *et al.*, 1993). While the phenotype of *snt-1* mutations is consistent with disruption of vesicle fusion, additional data suggest that synaptotagmin may act in vesicle membrane retrieval. This is consistent with the observation that synaptic terminals in *snt-1* null mutants are highly depleted in vesicles, although vesicle components are properly localized (Jorgensen *et al.*, 1995).

1.8.9 Mutations affecting process placement

Many Unc mutants have abnormal neural processes, resulting from errors in axon outgrowth and/or guidance. The phenotype of such mutants can often be correlated with defective uptake of fluorescent dyes into sensory neurons that, because of the Unc phenotype, implies similar structural defects in motor neurons (e.g. *unc-33*, *unc-44*, *unc-51*, *unc-76*, and *unc-6*; Hedgecock *et al.*, 1985). Alternatively, abnormal processes can be detected by any method that allows visualization of the nervous system. For example, commissures can be visualized by indirect immunofluorescence using antibodies to the neurotransmitter GABA (e.g. *unc-6*, *unc-33*; *unc-51*, *unc-76*; McIntire *et al.*, 1992). This aspect of nervous system development is of particular importance to the study of *unc-119*, as the mutant phenotype appears to result from a disruption in axon outgrowth or guidance (see Chapter 4). The next sections describe the principles underlying axon guidance and outgrowth, with particular focus on those factors identified in *C. elegans*.

1.9 Axon guidance

1.9.1 Basic forces in axon guidance

The intricate pattern of neural connectivity in a nervous system relies on the accurate navigation of axons to their targets during development. Over a century ago, Ramón y Cajal described the specialized ending of an outgrowing axon, the growth cone, and proposed that neurons are attracted by diffusible factors secreted by their target (Ramón y Cajal, 1890, 1892). The term axon guidance, then, refers to the process by which migrating growth cones choose a particular path along which to lay down axons.

Presently, it is thought that migration of growth cones occurs by the reception of external guidance cues of two types, which each may be attractive (permissive) or repulsive (non-permissive; reviewed by Tessier-Lavigne and Goodman, 1996). Figure 1.14 shows the types of forces that may act on growth cones. Long-range cues are provided by diffusible factors (chemoattractive or chemorepulsive); short-range cues are received via cell-cell or cell-substrate contact (permissive or non-permissive). A surprising finding has been that growth cones often respond to redundant signals, as these guidance forces may act simultaneously. There is, however, a general tendency to make use of particular types of signals over others. In early development, pioneering neurons migrate in an otherwise featureless environment guided by long-range cues, while later axons grow (fasciculate) along the pioneering neurons by following these same cues or by responding to new signals provided by the pioneering neurons (Tessier-Lavigne and Goodman, 1996).

Much of the current work on axon guidance has centred on the identification of signaling molecules and their receptors. Two model systems have been the focus of study for describing growth cones and chemoaffinity theory: The optic tectum, which illustrates the principle of topographic maps in vertebrates, and the ventral midline, which allows the study of guidance cues mediating outgrowth and fasciculation.

1.9.2 The Optic Tectum

The ganglionic cells of the vertebrate retina send processes to the optic tectum of the brain in an arrangement that preserves the topographic arrangement of photoreceptors (reviewed in Nicholls *et al.*, 1992; Tessier-

Lavigne and Goodman, 1996). During development, axons from temporally located cells innervate anteriorly located neurons, while axons from nasal ganglion cells innervate the posterior tectum (see Figure 1.15). Explant experiments in chick have shown that anterior tectal membranes will support the growth of temporal ganglionic neurites, but posterior tectal membranes repel temporal axons by inducing growth cone collapse (Walter *et al.*, 1987).

Recently, putative factors mediating this repulsion have been identified molecularly, which are members of the Eph-related group of receptor tyrosine kinases (Hirai *et al.*, 1987). Two Eph ligands, RAGS (repulsive axon guidance signal) and ELF-1, are expressed in overlapping anterior-to-posterior gradients across the entire tectum (Drescher *et al.*, 1995; Cheng and Flanagan, 1994, 1995). Consistent with these ligands acting in a repellent manner, retinal axons are repelled by RAGS *in vitro*, while the axons themselves show a similar gradient of sensitivity (Drescher *et al.*, 1995; Monschau *et al.*, 1997). Similarly, ectopic expression of ELF-1 steers temporal axons away to abnormal anterior positions (Nakamoto *et al.*, 1996).

The effect of Eph ligands on retinal axons is consistent with the proposed role of Eph receptors in axon guidance (Brambilla and Klein, 1995). Several receptors for the retinal Eph ligands have been identified (Cheng *et al.*, 1995; Monschau *et al.*, 1997). One of these, Mek4, is expressed in a reciprocal gradient across the retina, with weakest expression occurring in nasal axons. The contribution of these receptors to retinotectal projection is still not known (Tessier-Lavigne and Goodman, 1996). The number of known Eph ligands and receptors is growing, as other areas of the nervous system have been found to make use of similar gradients (reviewed in Nieto, 1996; Müller *et al.*, 1996).

1.9.3 The Ventral Midline

Long-range cues also affect axon migrations in the vertebrate spinal cord. Cell bodies of commissural interneurons located in the dorsal part of the developing cord first send processes circumferentially, then turn their axons toward the ventral midline as an intermediate target. These axons subsequently cross the midline and project longitudinally toward the brain (Nicholls *et al.*, 1992). The ventral floor plate promotes ventral outgrowth chemotropically (Tessier-Lavigne *et al.*, 1988). Initial studies suggested that neural cell adhesion molecules (N-CAMs) may play a direct role in guidance, as commissural axons express N-CAM throughout their length, along with the immunoglobulin/fibronectin type III (Ig/FNIII) superfamily member TAG-1 until they turn longitudinally, where TAG-1 expression is replaced by L1 (Figure 1.16; Dodd *et al.*, 1988). Interference with N-CAM interactions causes growing neurons to avoid the floor plate, consistent with requirement of N-CAM for contact-mediated interaction between commissural growth cones and the floor plate (Stoeckli *et al.*, 1997).

1.9.4 The netrin family of guidance molecules and their receptors

Long-range guidance of commissural axons toward the ventral midline has been found to rely on a conserved group of chemoattractants, the netrins. The *C. elegans* genes *unc-5*, *unc-6* and *unc-40* are involved in specifying guidance cues for the migration of mesodermal cells and pioneering axons (Hedgecock, 1990). Mutations in *unc-6* disrupt dorsal and ventral migrations, *unc-5* affects dorsal migrations, and *unc-40* mainly affects ventral migrations. Molecular identification of vertebrate and *Drosophila* homologues of these genes has demonstrated the existence of a conserved mechanism of axon outgrowth.

The gene product of *unc-6* is a 600-amino acid protein related to the extracellular matrix protein laminin, a basement membrane component that provides an adhesive matrix substrate for cells (Ishii *et al.*, 1992; Timpl and Brown, 1996). UNC-6 shows ~50% conservation with the vertebrate netrin family (Serafini *et al.*, 1994). Netrins were identified independently as molecules expressed in the ventral floor plate, capable of attracting commissural growth cones (Figure 1.16; Kennedy *et al.*, 1994). In the developing chicken spinal cord, netrin-1 is expressed in floor plate cells, and netrin-2 is expressed in the ventral two-thirds of the spinal cord. Consistent with a similar role for UNC-6, expression of an epitope-tagged version of the gene is limited to pioneering neurons and neuroglia of the ventral midline (Wadsworth *et al.*, 1996). These data suggest that UNC-6/netrins establish a dorsal-to-ventral gradient that serves to attract or repel circumferentially migrating growth cones (Wadsworth and Hedgecock, 1996).

The functional similarity of these guidance cues is consistent with the anatomical similarities of the vertebrate spinal cord and the *C. elegans* nervous system: The ventral floor plate, derived from midline neural ectoderm cells, is homologous to the ventral nerve cord of other phyla (Hedgecock and Hall, 1990). It has since been shown that netrin-1 has a simultaneous, chemorepulsive effect on trochlear motor axons, that grow away from the floor plate (Colamarino and Tessier-Lavigne, 1995).

The product of the *unc-5* gene is a novel cell surface receptor (Leung-Hagesteijn *et al.*, 1992). UNC-5 is required for migrating cells to be oriented away from ventral netrin sources (Hedgecock *et al.*, 1990; Leung-Hagesteijn *et al.*, 1992). Furthermore, *unc-5* expression appears to be necessary and sufficient for the ability of growth cones to be repelled by UNC-6, as ectopic expression in touch neurons, which do not normally express *unc-5,* causes their axons to grow dorsally (Hamelin *et al.*, 1993). Recently, vertebrate homologues of UNC-5 have been identified; cells expressing either UNC5H1 or UNC5H2 allows netrin-1 to bind them with high affinity (Leonardo *et al.*, 1997), suggesting that UNC-5 and UNC-6 may interact directly.

The *unc-40* gene encodes a homologue of the DCC subfamily of the Ig superfamily of cell surface receptors, initially thought to be inactivated in human colorectal cancer (Chan *et al.*, 1996; Hedrick *et al.*, 1994; Fazeli *et al.*, 1997). A conserved role for DCC in axon guidance is supported by studies in insects and mouse: Mutation in a *Drosophila* homologue, *frazzled*, results in commissural guidance defects similar to loss of netrin sources (Kolodziej *et al.*, 1996), and loss of *Dcc* in mouse results in defects in commissural axon guidance and brain development (Fazeli *et al.*, 1997). As was shown for UNC-5, transgene expression of DCC allows netrin-1 to bind cells directly, suggesting that netrins may directly bind DCC/UNC-40 (Keino-Masu *et al.*, 1996). However, a precise role for UNC-40 in attractive netrin-based motility is not yet clear, as neurons that express *unc-5*, which send axons away from netrin sources, also express *unc-40* (Chan *et al.*, 1996). As dorsal migrations are not strongly affected in *unc-40* mutants (Hedgecock *et al.*, 1990), UNC-40 may be playing an accessory or redundant role in the repulsive response.

A diagram of the pattern of expression of netrin receptors in *C. elegans*, rat spinal cord and *Drosophila* is given in Figure 1.17, and a model for netrin-based growth cone motility is shown in Figure 1.18.

1.9.5 Genes affecting axon guidance in fascicles

Cell adhesion molecules (CAMs) of the immunoglobulin and cadherin superfamilies, receptor protein tyrosine kinases (RPTKs), and receptor protein tyrosine phosphatases (RPTPs) are the main types of factors involved in growth of axons in fascicles (Tessier-Lavigne and Goodman, 1996). Both fasciculation and defasciculation can be regulated, for example, by expression of different combinations of CAMs (reviewed in Chiba and Keshishian, 1996).

Recently, mutations causing guidance defects in the *C. elegans* ventral nerve cord (VNC) have been described (Wightman *et al.*, 1997). Mutation of the genes *unc-3* and *unc-30* affect the growth of pioneering axons in the VNC, while mutations in *unc-42*, *unc-115*, *enu-1*, and *fax-1* affect the ability of later neurons to migrate along the pioneering axons. The UNC-30 gene product is a homeodomain protein that likely regulates expression of other genes (Jin *et al.*, 1994). The other genes in this group have not been identified molecularly, and the conservation of proteins involved specifically in *C. elegans* fasciculation has yet to be demonstrated (Wightman *et al.*, 1997).

1.10 Growth cone motility and axonogenesis

1.10.1 The study of axonogenesis

For axon guidance to occur, reception of guidance cues must be manifested as changes in cell shape; specifically, the growth cone must advance the growing axon in the direction established by external signals. Compared with axon guidance, less is known genetically about factors mediating growth cone motility and axonogenesis. In recent years, the molecular identification of particular Unc genes from *C. elegans* has expanded the use of the nematode for genetic analysis of factors required for axonogenesis. Many of these factors have been conserved, and in the general discussion (Chapter 6) a conserved role for UNC-119 in axonogenesis is proposed.

1.10.2 Structure of the growth cone

Ramòn y Cajal was the first to describe the "growth cone," a specialized structure present at the tip of growing axons (Ramòn y Cajal, 1890). This structure explores the environment and allows extension of the axon behind it in the appropriate direction. Therefore, the growth cone couples extracellular cues with change in cell shape during morphogenesis.

Migrating growth cones continually extend and retract two types of structures from the cell surface, filopodia and lamellipodia (Tanaka and Sabry, 1995). Filopodia are thin spike-like projections comprised of a dense, cross-linked bundle of actin filaments (Figure 1.19). Lamellipodia, which often form between filopodia, are broad cytoplasmic extensions that consist of long criss-crossed actin filaments. Lamellipodia and filopodia appear to allow detection of the extracellular environment (sensitive to minute changes in growth factor concentrations), and mediate substrate adhesion.

While the growth cone interior is composed primarily of actin, the interior of the formed axon consists of microtubules (Nicholls *et al.*, 1992). The ordered structure of the axoneme will later serve as the conduit along which intracellular vesicles are transported (see §1.8.7).

1.10.3 Three stages of growth cone motility

The process by which extracellular cues, received by filopodia and lamellipodia, result in axon formation, occurs in three steps: exploration, site selection, and site stabilization. Growth cone exploration involves the extension and retraction of lamellipodia and filopodia in three dimensions. This process depends on the local polymerization and depolymerization of actin (Stossel, 1993), and the translocation of actin filaments away from the leading edge of the growth cone (reviewed in Nicholls *et al.*, 1992). Binding of a cell surface receptor to a ligand is proposed to recruit a complex that physically couples the receptor to the actin filaments, preventing their flow. The lengthening of the actin filaments, by continual addition of actin monomers, results in cell protrusion (Figure 1.20). Proteins of the Rho subfamily of Ras-related GTPases are candidates for such transducers (Ridley and Hall, 1992), and recently a new Rho family member has been implicated in axon guidance and cell migration in *C. elegans* (Zipkin *et al.*, 1997).

Site selection involves two major changes in a region chosen for future growth: accumulation of actin polymer, followed by invasion of microtubules. Stable contact of a filopodium or lamellopodium with a favourable substrate causes local accumulation of actin, presumably by a similar mechanism as with cell protrusion. Microtubules, extending from the axon, then invade the region of actin accumulation. This may be because loss of actin from the lamella is more permissive to microtubule growth, or because microtubule polymerization is selectively stabilized by factors accumulating at the contact site (Tanaka and Sabry, 1995).

The last stage of growth cone migration is site stabilization. Microtubules, previously in a state of dynamic flux, become organized to form an ordered parallel bundle. The membrane of the growth cone collapses around this bundle to form the axon tube (Figure 1.21). The stability of axonal microtubules results from the binding of microtubule-associated proteins, or MAPs (Dreschel *et al.*, 1992; Kaech *et al.*, 1996).

1.10.4 Genes affecting axonogenesis in C. elegans

Mutations disrupting axon outgrowth would be expected to cause an Unc phenotype in nematodes. Specifically, loss of a factor required for proper growth cone migration, or stabilization of the newly formed axoneme, would be expect to result in severe structural defects in neurons. In *C. elegans* axonogenesis, mutants can be grouped based on whether they affect commissural neurite growth, fasciculation in process bundles, or both (McIntire *et al.*, 1992). Genes in the first class (*unc-5*, *unc-6* and *unc-40*) are involved in the production and reception of the netrin guidance cue itself, and therefore do not act in axon formation *per se* (see §1.9.4).

The genes *unc-34*, *unc-71* and *unc-76* are involved in fasciculation of axons in the ventral nerve cord (McIntire *et al.*, 1992). In mutants, the VNC is often split into multiple bundles, and some longitudinal axons terminate prematurely. The defects probably result from alterations in fasciculative signals (Antebi *et al.*, 1996). Recently, the gene product of *unc-76* has been found to define a new family of proteins, the FEZ family, with homologues in rat and human (Laird and Bloom, 1997).

Mutations in *unc-14*, *unc-33*, *unc-44*, *unc-51* and *unc-73* cause multiple defects in axonogenesis, including structural abnormalities, extra branches and abnormal fasciculation (Hedgecock *et al.*, 1985; McIntire *et al.*, 1992). This suggests that the gene products of these loci are structural or regulatory proteins affecting formation of the axonal cytoskeleton, a prediction that has been borne out by recent molecular data (Antebi *et al.*, 1996).

Collapsin, a member of the semaphorin family, can inhibit growth cone extension during axon outgrowth (Muller *et al.*, 1996). The UNC-33 protein is homologous to the collapsin-response mediator protein CRMP-62 identified in rat (Goshima *et al.*, 1995), and may therefore be a regulator of growth cone collapse. Consistent with a role in the cytoskeleton, *unc-33* mutants have excess axonal microtubules in sensory neurons (Hedgecock *et al.*, 1985). The product of the *unc-44* gene is related to ankyrins (Otsuka *et al.*, 1995), cytoskeletal adaptors that couple spectrin and actin filaments to membrane proteins. A direct role in the cytoskeleton is therefore implied for UNC-44.

The UNC-51 protein is a novel serine/threonine kinase expressed throughout the nervous system, and that may therefore be involved in signal transduction (Ogura *et al.*, 1994). The expression of the *unc-14* gene overlaps that of *unc-51*, and UNC-14 and UNC-51 are able to interact *in vitro* (Ogura *et al.*, 1997). This suggests that UNC-14 may be a regulator of UNC-51 activity (Ogura *et al.*, 1997).

In summary, the molecular identification of *C. elegans* genes affecting axon outgrowth and fasciculation, coupled with the surprising finding that most of these factors are conserved, have justified the use of the nematode as a model system for neurobiological study. A more complete understanding of the conserved mechanisms at work in neural development in all animals will certainly emerge as more neural genes are characterized.

1.11 Background to the thesis

The stage has been set, therefore, for the use of *C. elegans* as a system in which to study neural development genes. At the outset of this thesis, the field of *C. elegans* molecular neurobiology was just beginning to identify and characterize genes with a role in axon outgrowth and axon guidance. The molecular identification of these genes, and the elucidation of interactions among the gene products, are now the subject of ongoing investigation.

Prior to the start of this thesis, the previous work on the *unc-119* gene was only preliminary. The *unc-119* phenotype was characterized by strong body paralysis, scorable at hatching, that persisted through adulthood. The earliest strain now known to carry a lesion in *unc-119* is CB1517, bearing the duplication-deficiency *eDf2*; *eDp6 III*. This strain harbours a mutually complementary set of linkage group *III* chromosomal aberrations that were derived concomitantly following acetaldehyde mutagenesis (Hodgkin, 1980). This strain is homozygous for these chromosome *III* fragments and has an Unc phenotype, that was assumed to be a result of an interruption in a gene, or group of genes, at the common breakpoint of the two lesions. It was recognized as carrying a defect in *unc-119* following the recovery of the spontaneous mutation *e2498* by D. Pilgrim (see Chapter 2). The similarity between the *e2498* and *eDf2*; *eDp6* phenotypes was pointed out by J. Hodgkin; a complementation test showed that *unc-119* (*e2498*) could not be complemented by either *eDf2* or *eDp6* (D. Pilgrim, pers. comm.). Figure 1.22 shows the chromosomal relationship between *unc-119* and the *eDf2* and *eDp6* aberrations. At this point, no further genetic or phenotypic characterization of *unc-119* or *eDf2*; *eDp6* was performed.

The common breakpoint of *eDf2* and *eDp6*, approximately 1/3 of the way from the right end of chromosome *III*, was used by *C. elegans* investigators as a convenient mapping tool for genes in the region. Loci on LG *III* could be complemented by either *eDf2* or *eDp6* (Hodgkin, 1980). These mapped genes included the cloned loci *tra-1*, the leftmost cloned locus on *eDp6* (Zarkower and Hodgkin, 1992) and the β-tubulin gene *tbb-1*, the farthest right on the *eDf2* chromosome (Sulston *et al.* 1992). As *unc-119* failed to be complemented by either *eDf2* or *eDp6*, the gene was localized to their common breakpoint. This lay between *tra-1* and *tbb-1*, in a ~1 Mb region spanning several overlapping YACs (yeast artificial chromosomes), that included many gaps lacking cosmid coverage. A few representative cosmids and YACs had been obtained from the *C. elegans* physical mapping project (Coulson *et al.*, 1986), although a detailed molecular characterization of the region had not yet been performed. Figure 1.22 shows a portion of the *C. elegans* physical map around the *unc-119* region.

1.12 Goals of this thesis

The aim of this thesis was to clone and characterize the *unc-119* gene of *C. elegans*, to determine its pattern of expression, and to gain an understanding of its role in nervous system development. Specific objectives were to:

1) Obtain new mutations in *unc-119* to establish a monogenic basis for the *unc-119* phenotype.

2) Localize the *unc-119* locus on the *C. elegans* physical map, and clone the entire gene by germline transformation rescue.

3) Determine the expression pattern of *unc-119* by using reporter gene fusions.

4) Characterize the behavioural and cellular aspects of the *unc-119* mutant phenotype.

5) Identify putative homologues of UNC-119 in other species, and test their function in *C. elegans*.

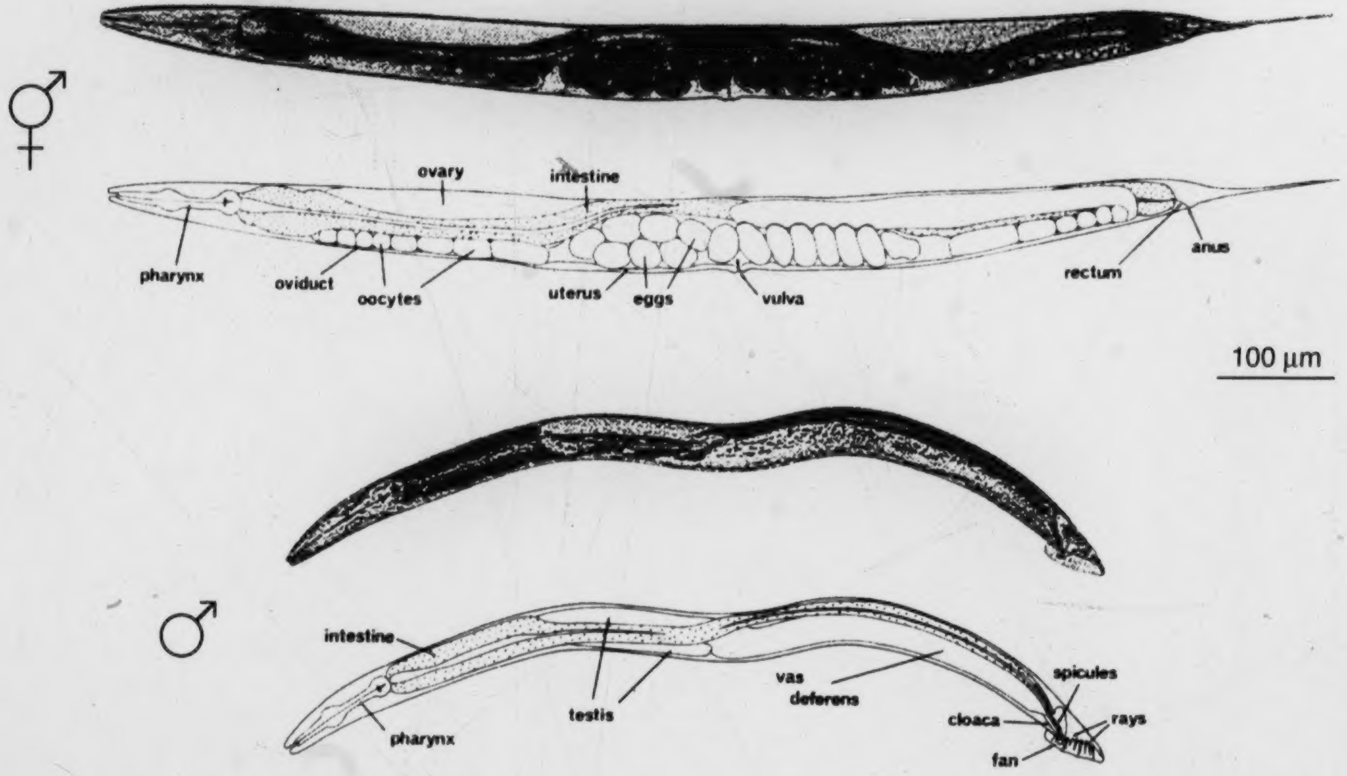


Figure 1.1 Appearance of wild type *Caenorhabditis elegans.* Top, light micrograph and line drawing of adult hermaphrodite. Bottom, micrograph and drawing of adult male. Anterior is to the left, and dorsal is at top. Figure reproduced, with permission, from Sulston and Horvitz (1977).

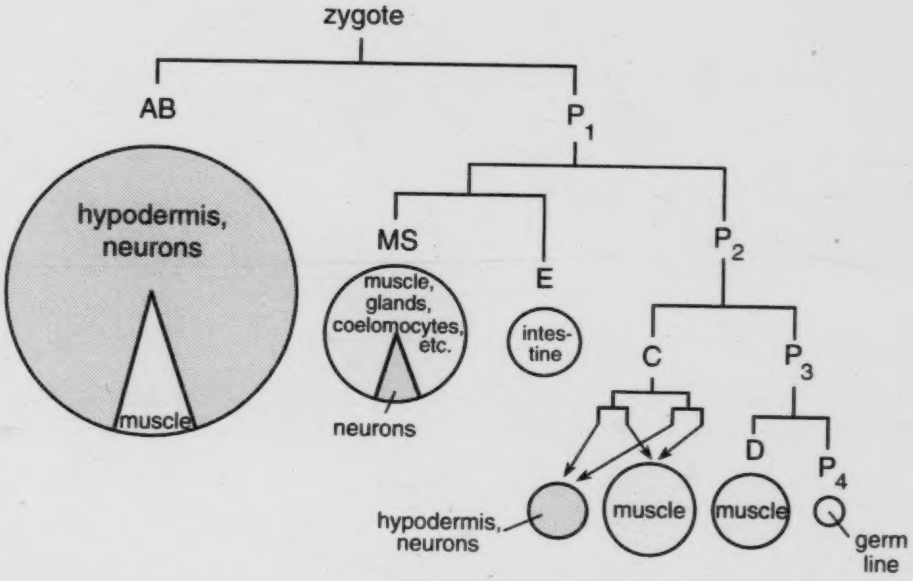


Figure 1.2 Generation of the six *C. elegans* founder cells, showing the cell types that are derived from them. Each horizontal line represents a cell division. The areas of the circles are approximately proportional to the numbers of cells in the adult that are derived from each founder. Figure redrawn from Sulston *et al.* (1983).

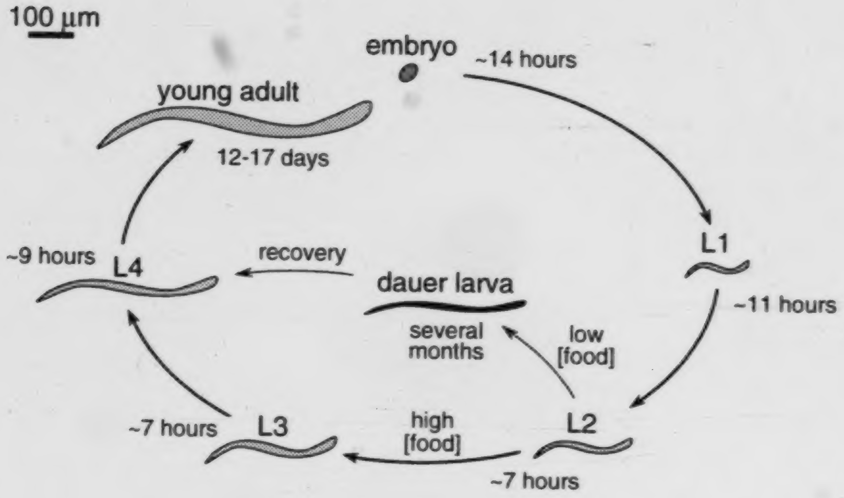


Figure 1.3 The life cycle of *C. elegans*. The approximate sizes and duration of each stage are shown. Animals hatch as L1 larvae. Progression through each subsequent larval stage is accompanied by a larval molt. Under conditions of starvation and crowding, L2 animals can develop into dauer larvae, which can persist for several months. Data from Wood (1988).

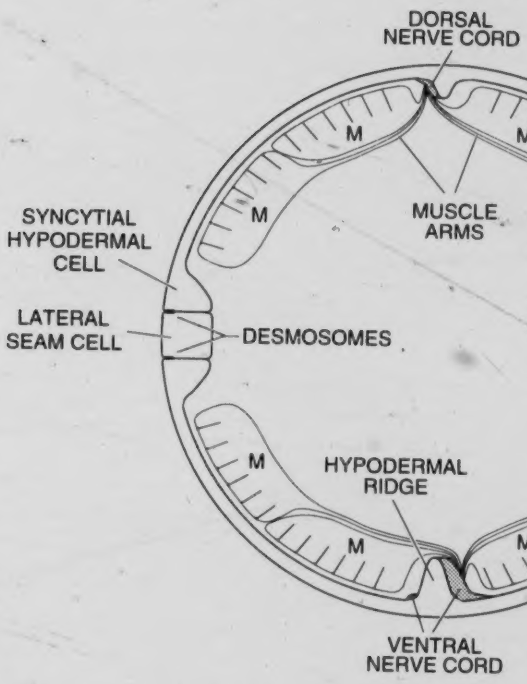


Figure 1.4 Diagram of a transverse section of the nematode at the midbody, showing the four muscle quadrants. The body tube consists of the syncytial cell and the lateral seam cells, joined by desmosomes. The dorsal and ventral nerve cords each run along a longitudinal hypodermal ridge. The four quadrants of body wall muscle are each composed of two rows of muscle cells (M), which send out processes (muscle arms) to the nerve cords to form neuromuscular junctions. Figure redrawn from White *et al.* (1986).

C. elegans Axon Scaffold

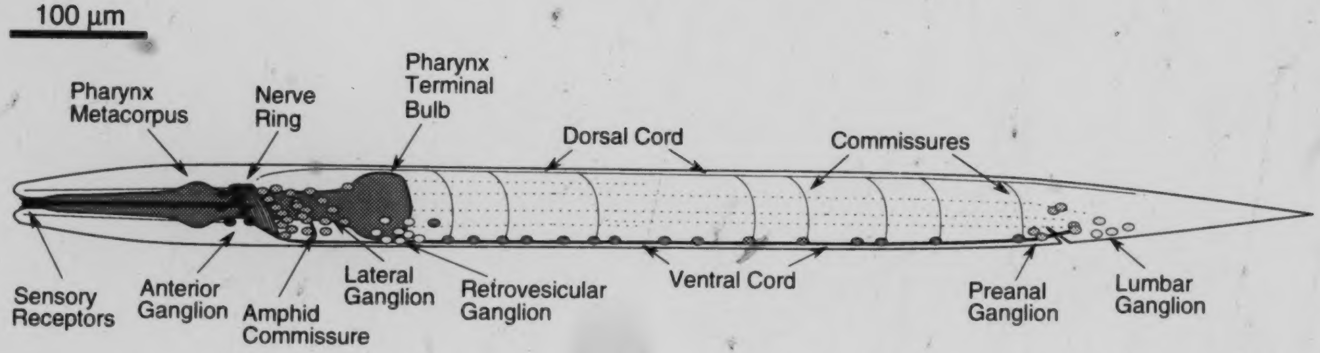


Figure 1.5 Schematic diagram of *C. elegans* showing the axon scaffold and ganglionic clusters. Representative cell bodies are shown as small circles. Dotted lines represent lateral processes. Anterior is to the left. Figure redrawn from Wadsworth and Hedgecock (1996).

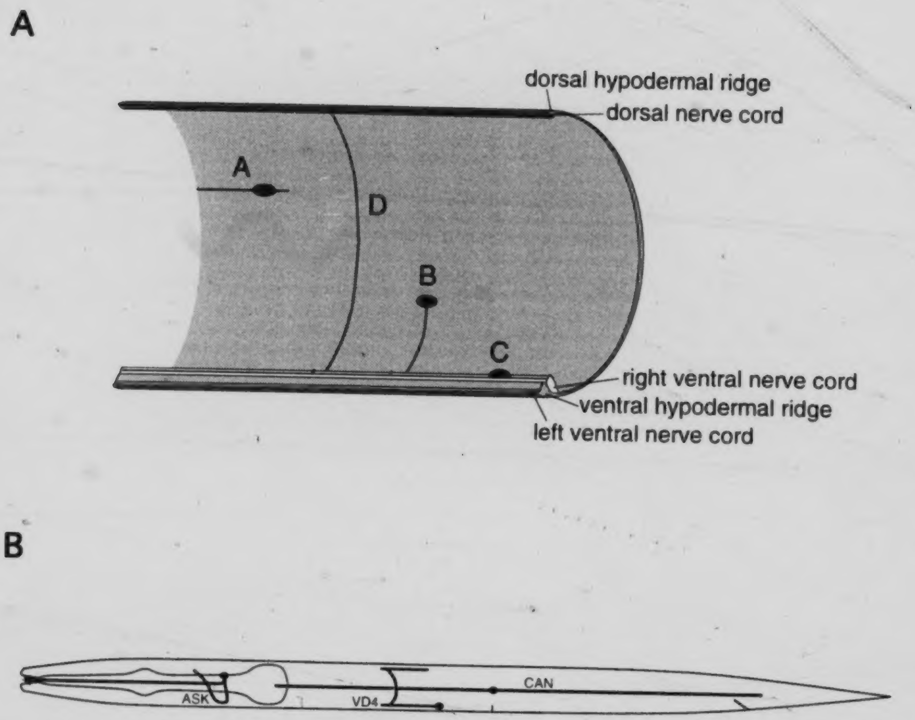


Figure 1.6 Axon morphology of *C. elegans* neurons. **A**, Lateral cut-away view showing axonal morphology of some neuron types: a lateral cell body and process (A) such as the neuron ALM; a lateral cell body (B) which extends a process ventrally to the ventral nerve cord, such as AVM; and a motor neuron, with cell body (C) adjacent to the ventral nerve cord, which sends a process in the ventral nerve cord, which branches to produce a commissure (D) which extends dorsally and enters the dorsal nerve cord. Figure redrawn and modified from McIntire *et al.* (1992). **B**, Axon morphology for three neurons. For ASK and CAN, only the left members of each pair are shown. Cell bodies are indicated by filled black ovals. Data compiled from White *et al.* (1988).

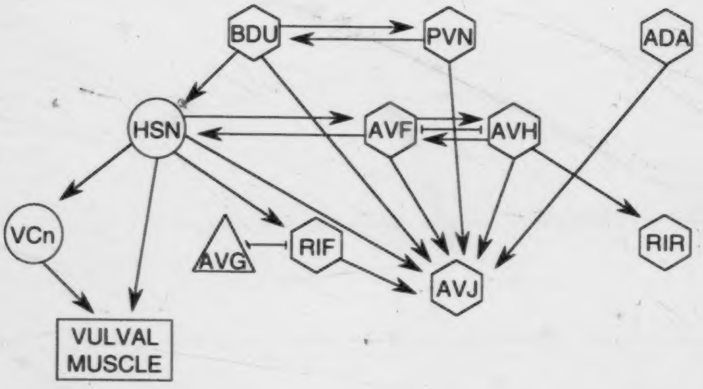


Figure 1.7 Neural connectivity diagram of the egg laying circuitry showing chemical synapses (arrows) and gap junctions (Ts). Neuron classes are indicated by symbol: triangle, sensory neuron; hexagon, interneuron; circle, motor neuron. Figure is redrawn from White *et al.* (1988).

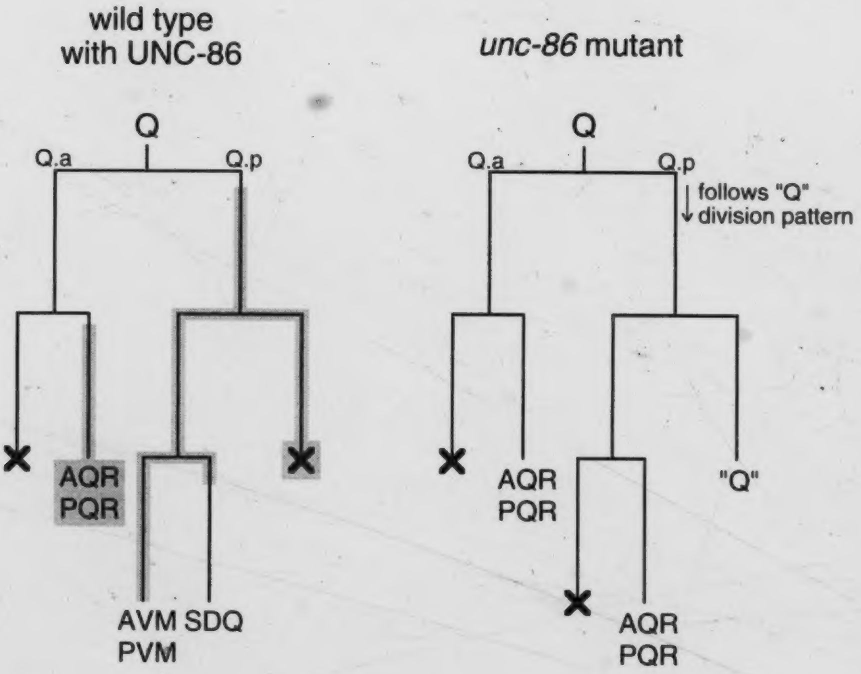
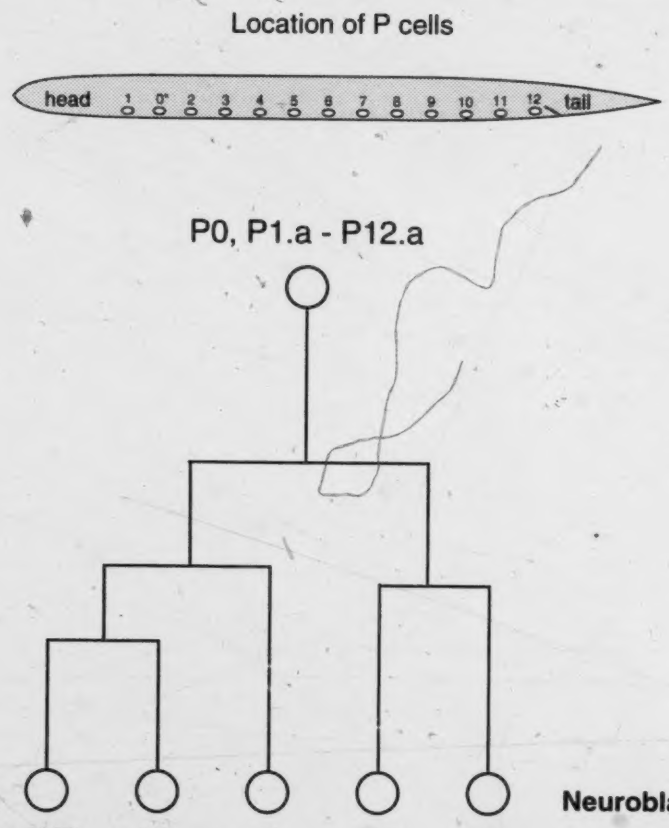


Figure 1.8 Cell lineage of the Q neuroblast in wild type and *unc-86* mutants. In the wild type, the neurons AVM, PVM and SDQ are generated from Q.p (posterior daughter of Q). In an *unc-86* mutant, these neurons are not made; instead, Q.p reiterates the division pattern of its mother cell Q, producing supernumerary AQR and PQR neurons. (X=programmed cell death; cells expressing *unc-86* are shown as shaded.) Figure redrawn from Ruvkun (1997).



Neuron Type					Neuroblast
AVF	VB	X	VA	VD	P0
AVF	VB	X	AS	VD	P1.a
VA	VB	X	AS	VD	P2.a
VA	VB	VC	AS	VD	P3.a-P8.a
VA	VB	X	AS	VD	P9.a-P10.a
VA	X	X	AS	VD	P11.a
VA	X	X	PDB	VD	P12.a

Figure 1.9 The P cells are the neuroblasts (top) that generate the L1-derived motor neurons in the ventral nerve cord. The P0 cell (*) is analogous to the anterior daughters of P1-P12.. Each undergoes a similar pattern of divisions generating neurons of similar type (with some exceptions, such as the occurrence of cell deaths, shown as 'X'). Figure redrawn from Chalfie and White (1988) and Ruvkun (1997).

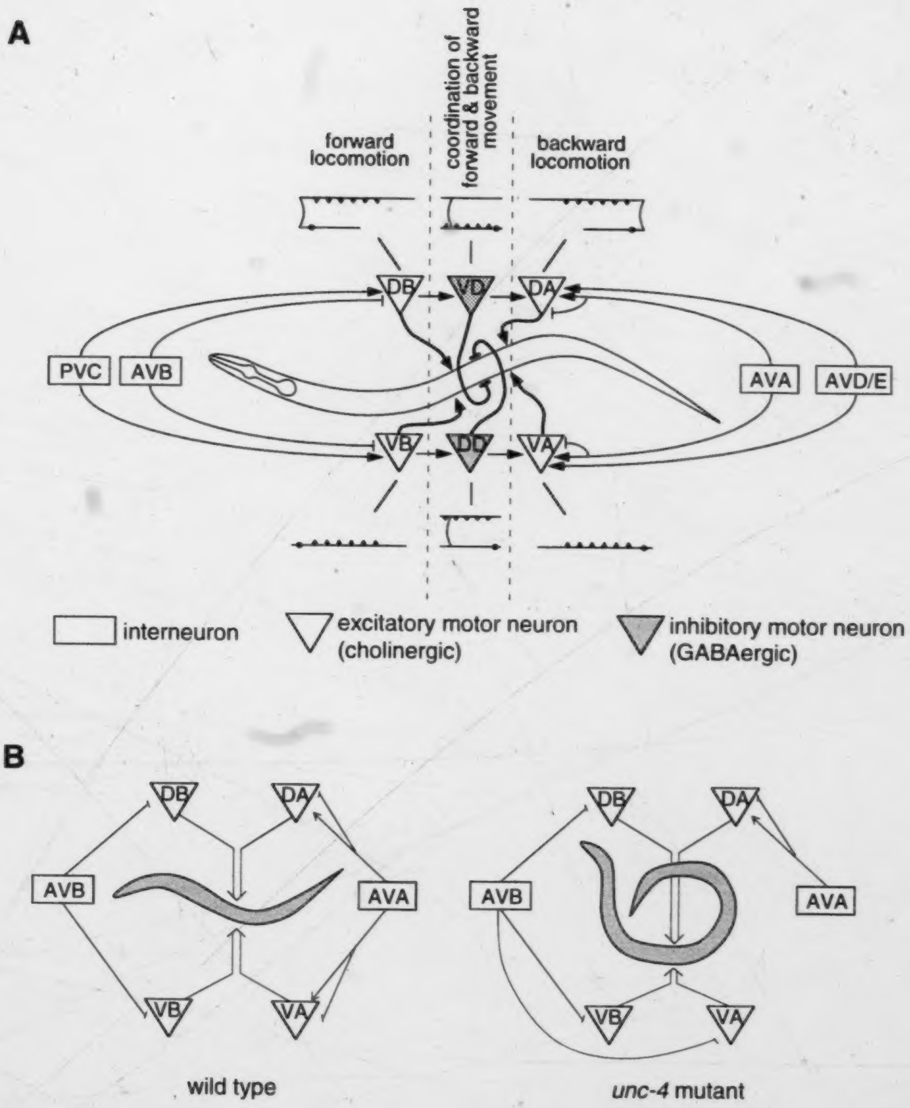


Figure 1.10 Circuitry involved in locomotion is disrupted in *unc-4*. **A**, Diagram showing representatives of each of the six classes of major motor neuron. The V neurons innervate ventral muscles; the D neurons innervate dorsal muscles. The morphologies of the neuron classes are shown. **B**, In *unc-4* mutants, the VA neurons have synaptic input characteristic of their VB sisters, resulting in defective backward locomotion. Symbols as in part A. Figures redrawn from White *et al.* (1992) and Driscoll and Kaplan (1997).

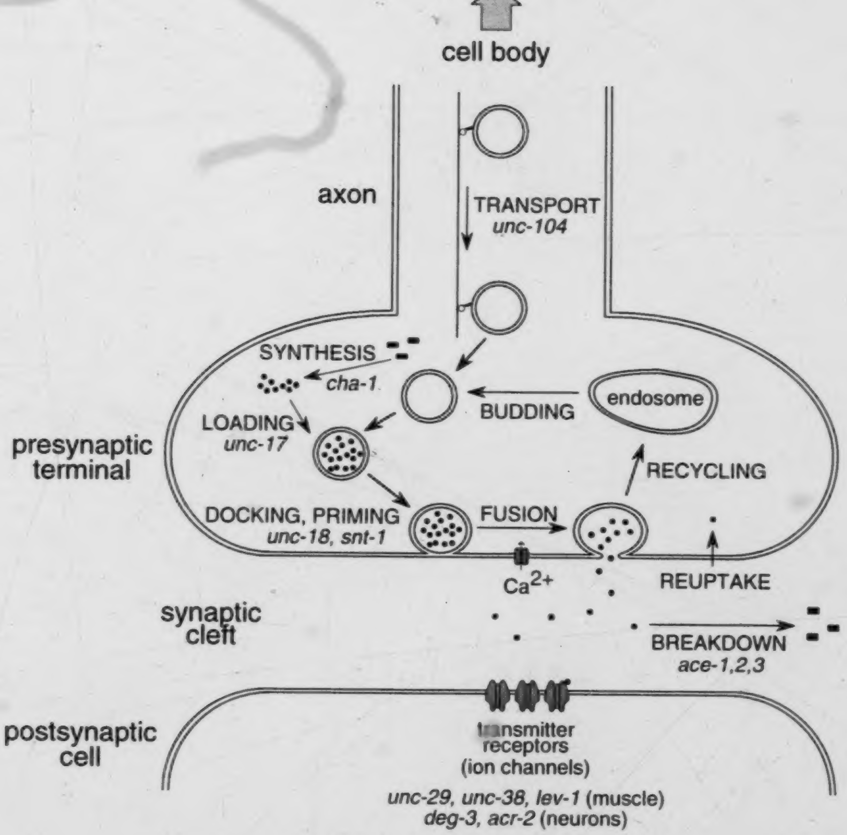


Figure 1.11 The steps in neurotransmission at a chemical synapse. Synaptic vesicle precursors are made in the cell body and transported to the presynaptic terminal, where neurotransmitter is synthesized from cell metabolites. The vesicles then are targeted to the membrane (docking) where they are primed for fusion. Calcium influx initiates fusion with the plasma membrane. Following transmitter release, vesicles are recycled by endosomal fusion. After the neurotransmitter acts on receptors on the postsynaptic cell, it is either reclaimed by an uptake transporter, or broken down enzymatically. Genes whose products function at particular steps in motor neurons are indicated. Components are not drawn to scale. Figure redrawn and modified from Südhof (1995) and Rand and Nonet (1997).

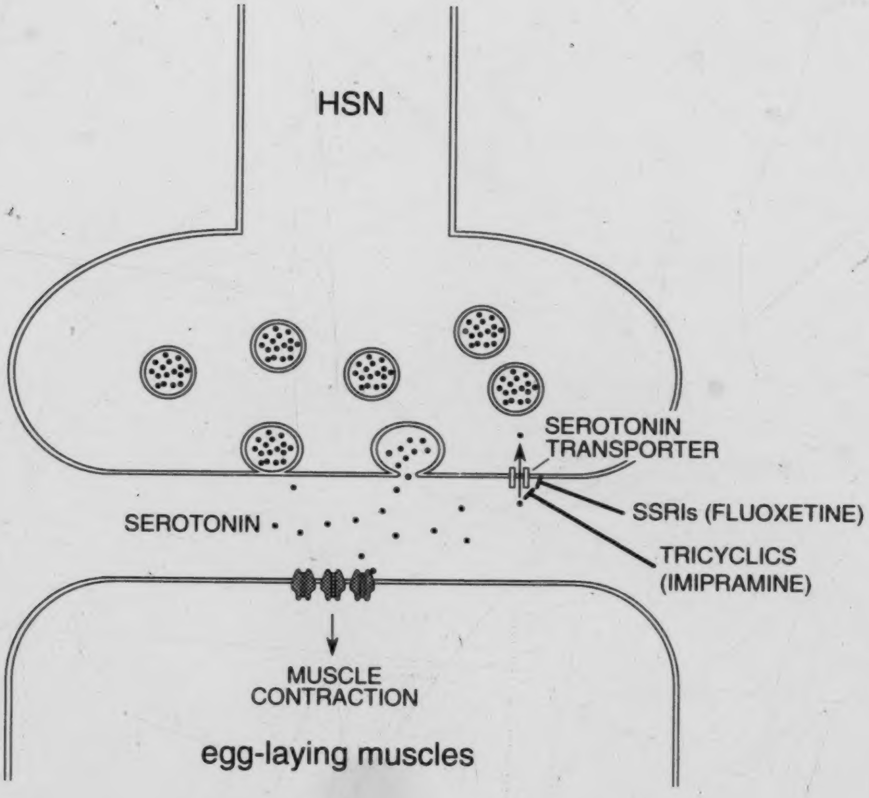


Figure 1.12 Model for egg laying, and the action of tricyclics and SSRIs. The HSNs release serotonin into the synaptic cleft, which stimulates receptors on the egg-laying muscles. These contract and allow the release of eggs from the uterus. Excess neurotransmitter is pumped back into the HSNs to prevent further stimulation. Exogenous serotonin overwhelms this reuptake system, and potentiates egg laying by directly stimulating the egg laying muscles, while fluoxetine and imipramine stimulate egg laying by preventing serotonin reuptake. ACh, which is required for egg laying, is not shown. Figure is modified from Weinshenker *et al.* (1995).

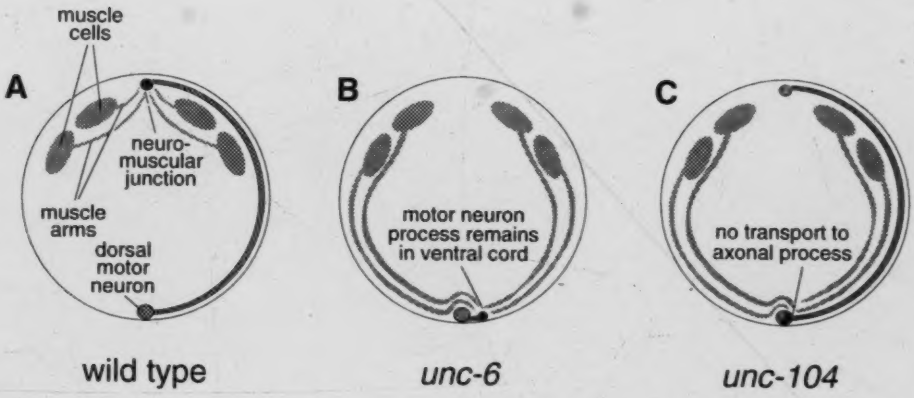


Figure 1.13 Transverse sections showing that muscle arms can grow out towards misplaced motor neurons, presumably by chemoattraction to substances released by synaptic vesicles. Dorsal is at top. **A**, In wild type, muscle arms grow toward the dorsal nerve cord. **B**, In an *unc-6* mutant, axons fail to extend dorsally, and synaptic vesicles accumulate at the shortened axon terminal, to which muscle arms extend. **C**, In *unc-104*, axons grow out correctly, but synaptic vesicles remain near the cell body. Figure redrawn from Jorgensen and Rankin (1997).

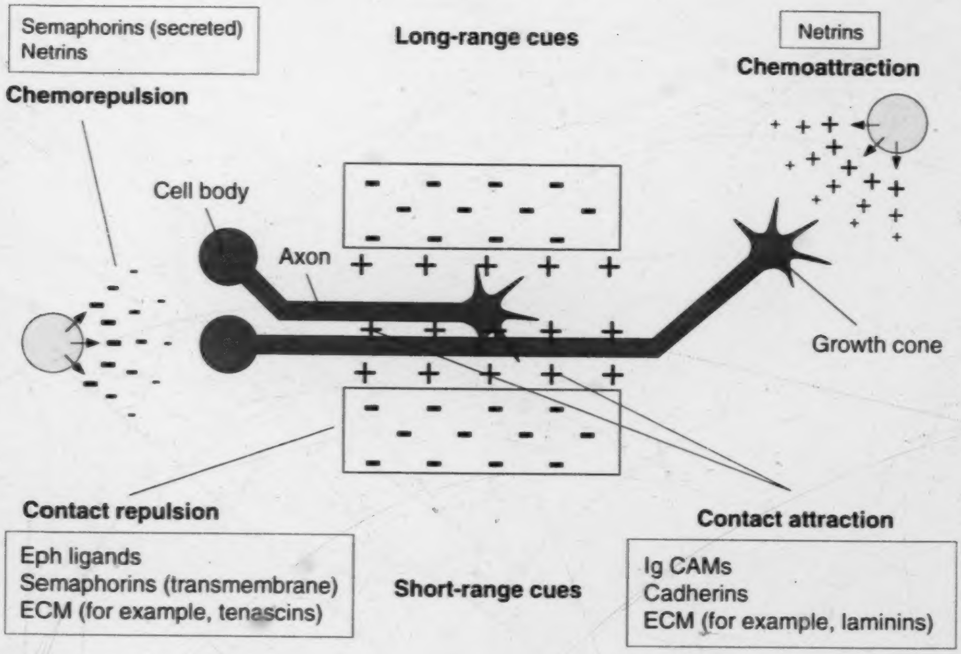


Figure 1.14 The main guidance forces that act during axon outgrowth. Long-range cues affect migration at a distance, as occurs in commissural outgrowth in *C. elegans*; these may be chemorepulsive or chemoattractive. Short-range cues affect outgrowth locally and are mediated by cell-cell (or cell-substrate) contact, as in the fasciculation of the ventral nerve cord. Short-range cues may also act repulsively or attractively. Figure redrawn and modified from Tessier-Lavigne and Goodman (1996).

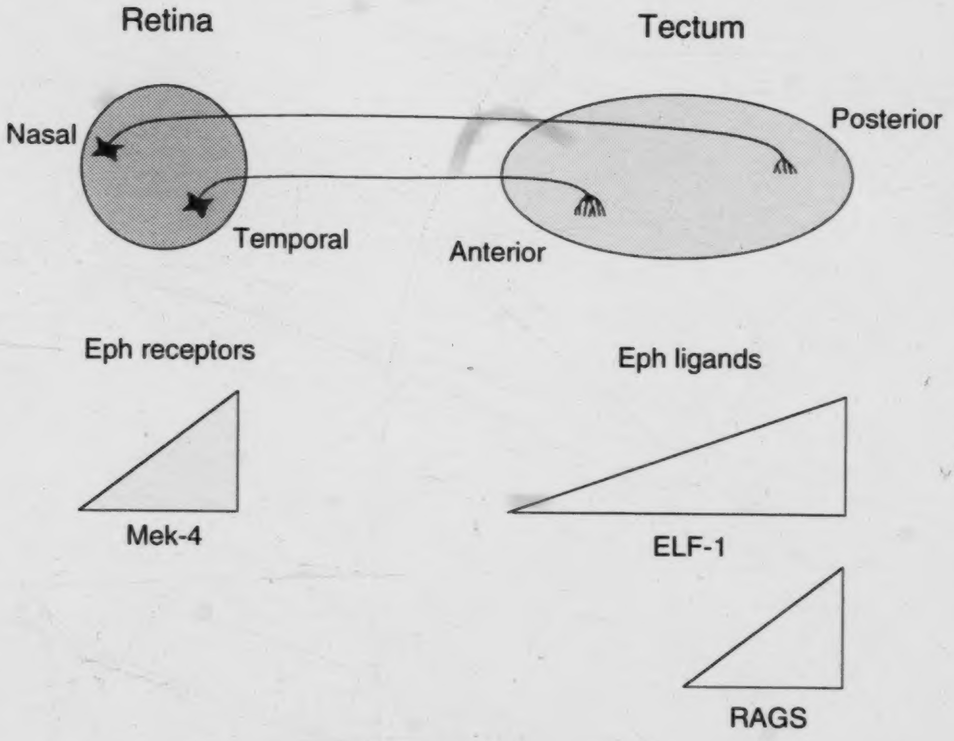


Figure 1.15 Retinotectal projection. Retinal ganglion cells in the chick embryo follow a precise spatial arrangement of innervation: Cells in the nasal retina innervate neurons in the posterior tectum, and cells in the temporal retina innervate neurons in the anterior tectum. The topographic map is specified by long-range gradient cues (see text). The change in height from left to right indicates a gradient of increasing concentration. Figure redrawn from Tessier-Lavigne and Goodman (1996).

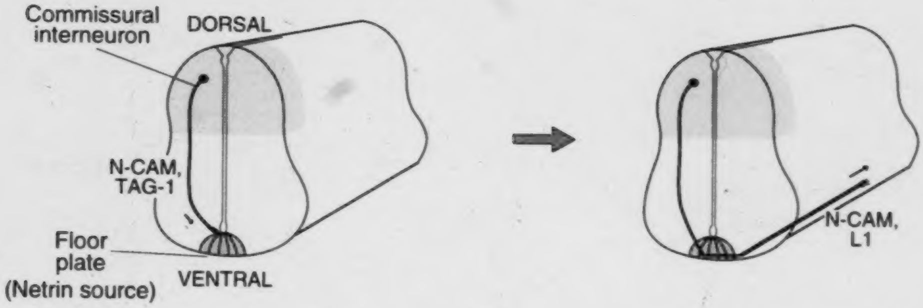


Figure 1.16 Growth of rat interneurons in the developing spinal cord. Dorsally located cell bodies send processes circumferentially toward the ventral floor plate, and then turn longitudinally. N-CAM is expressed along the entire length of the axon. TAG-1 (Axonin-1 in chick) is expressed until the longitudinal turn, where it is replaced by L1. Ventral growth depends upon chemoattraction to the floor plate, which expresses netrin. Figure redrawn and modified from Nicholls *et al.* (1992).

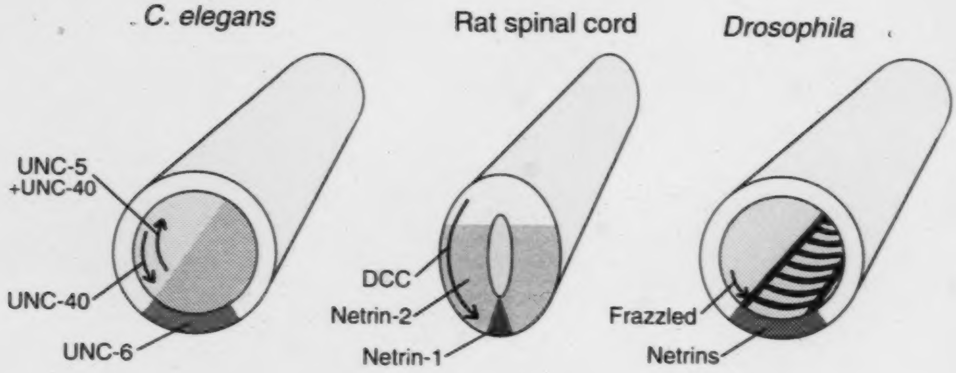


Figure 1.17 Cross-sections through *C. elegans*, rat spinal cord and *Drosophila* at early stages of development. The expression pattern of netrin receptors on axonal trajectories, and the netrins expressed at the ventral midline, are indicated. Figure redrawn from Drescher (1996).

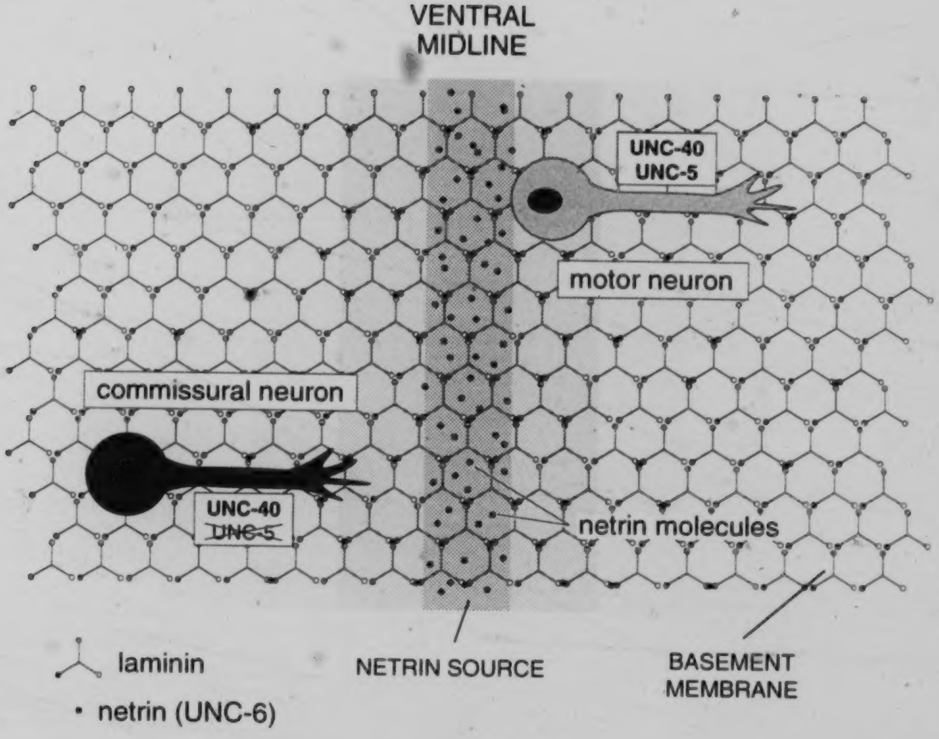


Figure 1.18 Model for netrin-based motility. Netrins secreted from midline cells could be tethered to basement membranes, forming a gradient. Commissural axons expressing UNC-40 orient growth cones toward the cue, while motor axons express both UNC-40 and UNC-5 and orient away from the midline. Figure modified from Hedgecock and Norris (1997).

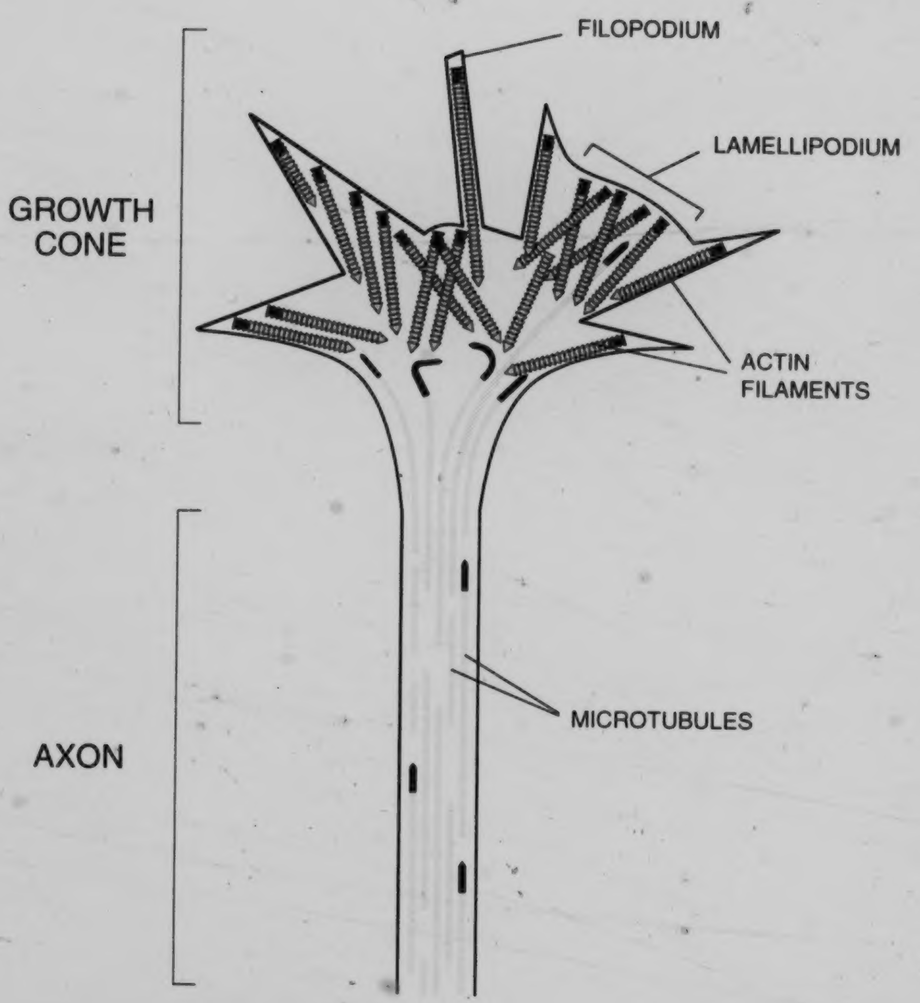


Figure 1.19 Schematic depiction of a growth cone, showing cytoplasmic projections and the distribution of actin and microtubules. Lamellipodia and filopodia protrude and retract rapidly to allow exploration of the growth cone microenvironment. The motility results from actin polymerization near the membrane (black segments), and the translocation of actin filaments rearward. Most microtubules are found in the axon; some extend into lamellipodia. Microtubules also change their shape dynamically by polymerization, movement and bending (black regions). Figure redrawn and modified from Tanaka and Sabry (1995).

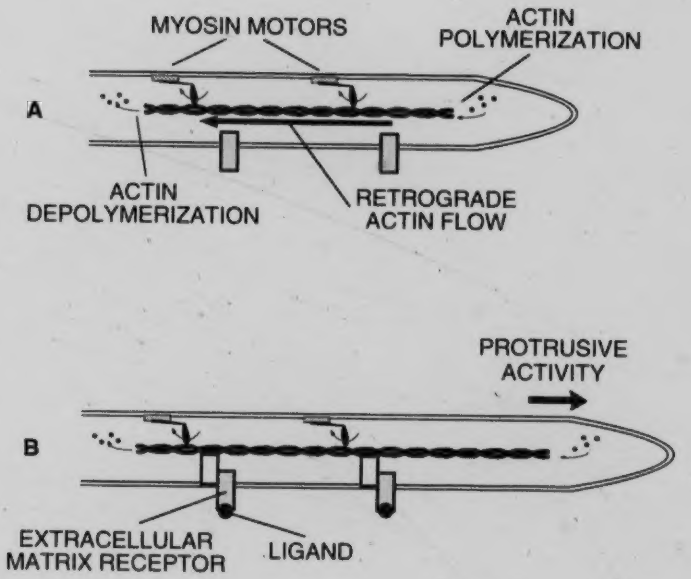


Figure 1.20 Model for growth cone motility. **A**, myosin-type motors drive retrograde flow of the actin filament. Polymerization occurs at the leading edge, while depolymerization occurs towards the lamella. **B**, protrusion of the growth cone occurs when a ligand binds an extracellular matrix receptor (shaded boxes), resulting in the binding of a protein complex (white boxes) with actin. This retards the retrograde flow, leading to protrusion of the membrane. Figure adapted from Nicholls *et al.* (1992) and Tanaka and Sabry (1995).

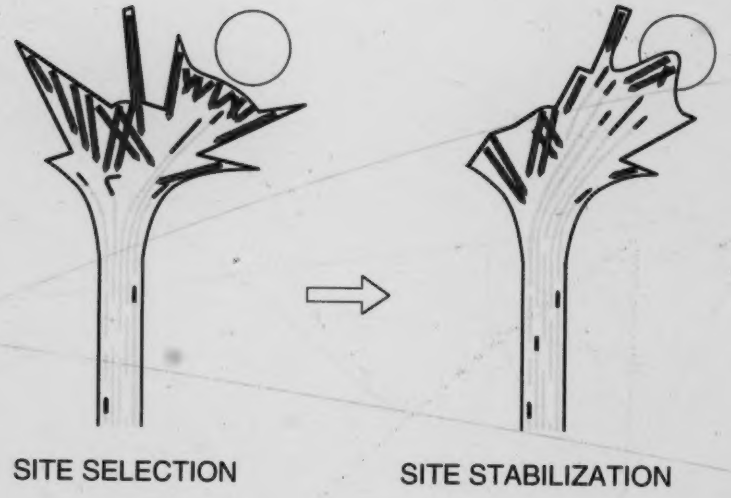


Figure 1.21 Actin accumulates where the growth cone meets the guidance cue (large circle), and is depleted in adjacent zones, during site selection. During site stabilization, microtubules invade the site and begin to form a bundle. The growth cone membrane collapses around this bundle, extending the axon, and advancing the leading edge of the growth cone. Figure redrawn from Tanaka and Sabry (1995).

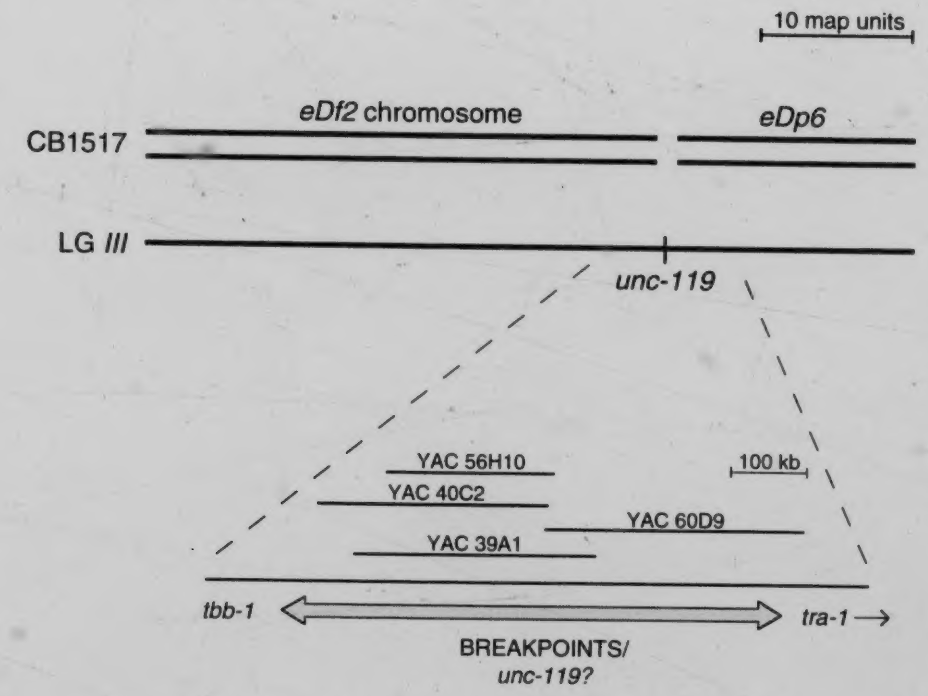


Figure 1.22 Diagram showing placement of *unc-119* at the breakpoints of the lesions *eDf2* and *eDp6*. The strain CB1517, which is homozygous for both *eDf2* and *eDp6*, has an Unc-119 phenotype. The portion of LG*III* remaining in *eDf2* is indicated as '*eDf2* chromosome.' A segment of the physical map is shown below the genetic map, indicating the ~1 megabase region that contains *unc-119* and the breakpoints of *eDf2* and *eDp6*. Some representative YACs (yeast artificial chromosomes) are shown.

1.11 Bibliography

Alfonso, A., Grundahl, K., Duerr, J. S., Han, J. P. and Rand, J. B. (1993). The *Caenorhabditis elegans unc-17* gene: A putative vesicular acetylcholine transporter. Science *261*, 617-619.

Alfonso, A., Grundahl, K., McManus, J. R., Asbury, J. M. and Rand, J. B. (1994). Alternative splicing leads to two cholinergic proteins in *Caenorhabditis elegans*. J. Mol. Biol. *241*, 627-630.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. NAR *25*, 3389-3402.

Antebi, A., Norris, C. R., Hedgecock, E. M., and Garriga, G. (1996). Cell and Growth Cone Migrations. in *C. elegans II.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Avery, L. (1993). The genetics of feeding in *Caenorhabditis elegans.* Genetics *133*, 897-917.

Avery, L. and Horvitz, H. R. (1989). Pharyngeal pumping continues after laser killing of the pharyngeal nervous system of *C. elegans.* Neuron *3*, 473-485.

Avery, L., Lockery, S., and Raizen, D. M. (1995). Electrophysiological methods in *C. elegans.* in *Caenorhabditis elegans*: *Modern Biological Analysis of an Organism* (eds Epstein, H. F. and Shakes, D. C.) Methods Cell Biol. *48*, 251-269.

Bargmann, C. I. and Horvitz, H. R. (1991a). Chemosensory neurons with overlapping functions direct chemotaxis to multiple chemicals in *C. elegans.* Neuron *7*, 729-742.

Bargmann, C. I. and Horvitz, H. R. (1991b). Control of larval development by chemosensory neurons in *Caenorhabditis elegans.* Science *251*, 1243-1246.

Bloom, L. and Horvitz, H. R. (1997). The *Caenorhabditis elegans* gene *unc-76* and its human homologs define a new gene family involved in axonal outgrowth and fasciculation. PNAS(USA) *94*, 3414-3419.

Brambilla, R., and Klein, R. (1995). Telling axons where to grow: a role for Eph receptor tyrosine kinases in guidance. Mol Cell Neurosci *6*, 487-495.

Brenner, S. (1974). The genetics of *Caenorhabditis elegans.* Genetics *77*, 71-94.

Burr, A. H. (1985). The photomovement of *Caenorhabditis elegans*, a nematode which lacks ocelli: Proof that the response is to light, not radiant heating. Photochem. Photobiol. *41*, 577-582.

Chalfie, M., Horvitz, H. R., and Sulston, J. E. (1981). Mutations that lead to reiterations in the cell lineages of *C. elegans.* Cell *24*, 59-69.

Chalfie, M. and White, J. (1988). The Nervous System. in *The Nematode Caenorhabditis elegans.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Chalfie, M,. Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. Science *263*, 802-805.

Chamberlin, H. M. and Sternberg, P. W. (1995). Mutations in the *Caenorhabditis elegans* gene *vab-3* reveal distinct roles in fate specification and unequal cytokinesis in an aymmetric cell division. Dev. Biol. *170*, 679-689.

Chan, S. S. -Y., Zheng, H., Su, M. -W., Wilk, R., Killeen, M. T., Hedgecock, E. M., and Culotti, J. G. (1996). UNC-40, a *C. elegans* homolog of DCC (Deleted in Colorectal Cancer), is required in motile cells responding to UNC-6 netrin cues. Cell *87*, 187-195.

Cheng, H. J. and Flanagan, J. G. (1994). Identification and cloning of ELF-1, a developmentally expressed ligand for the Mek4 and Sek receptor tyrosine kinases. Cell *79*, 157-168.

Cheng, H. J., Nakamoto, M., Bergemann, A. D., and Flanagan, J. G. (1995). Complementary gradients in expression and binding of ELF-1 and Mek4 in development of the topographic retinotectal projection map. Cell *82*, 371-381.

Chiba, A. and Keshishian, H. (1996). Neuronal Pathfinding and Recognition: Roles of Cell Adhesion Molecules. Dev. Biol. *180*, 424-432.

Colamarino, S. A. and Tessier-Lavigne, M. (1995). The axonal chemoattractant netrin-1 is also a chemorepellent for trochlear motor axons. Cell *81*, 621-629.

Coulson, A. R., Sulston, J., Brenner, S., and Karn, J. (1986). Toward a physical map of the genome of the nematode *C. elegans.* PNAS(USA) *83*, 7821-7825.

Coulson, A. R., Waterston, R. H., Kiff, J. E., Sulston, J. E., and Kohara, Y. (1988). Genome linking with yeast artificial chromosomes. Nature *335*,184-186.

Coulson, A. R., Kozono, Y., Lutterbach, B., Shownkeen, R., Sulston, J., and Waterston, R. (1991). YACs and the *C. elegans* genome. BioEssays *13*, 413-417.

Desai, C., Garriga, G., McIntire, S. L., and Horvitz, H. R. (1988). A genetic pathway for the development of the *Caenorhabditis elegans* HSN motor neurons. Nature *336*, 638-646.

Desai, C. and Horvitz, H. R. (1989). *Caenorhabditis elegans* Mutants Defective in the Functioning of the Motor Neurons Responsible for Egg Laying. Genetics *121*, 703-721.

Dodd, J., Morton, S. B., Karagogeos, D., Yamamoto, M., and Jessell, T. M. (1988). Spatial regulation of axonal glycoprotein expression on subsets of embryonic spinal neurons. Neuron *1*, 105-116.

Drescher, U. (1996). Netrins find their receptor. Nature *384*, 416-417.

Drescher, U., Kremoser, C., Handwerker, C., Löschinger, J., Noda, M., and Bonhoeffer, F. (1995). *In vitro* guidance of retinal ganglion cell axons by RAGS, a 25 kDa tectal protein related to ligands for Eph receptor tyrosine kinases. Cell *82*, 359-370.

Dreschel, D. N., Hyman, A. A., Cobb, M. H., and Kirschner, M. W. (1992). Modulation of the dynamic instability of tubulin assembly by the microtubule-associated protein tau. Mol. Biol. Cell. *3*, 1141-1154.

Driscoll, M. (1992). Molecular genetics of cell death in the nematode *Caenorhabditis elegans*. J. Neurobiol. *23*, 1327-1351.

Driscoll, M. and Kaplan, J. (1997). Mechanotransduction. in *C. elegans II*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Erickson, J., Varoqui, H., Shafer, M., Modi, W., Diebler, M., Weihe, E., Rand, J., Eiden, L., Bonner, T., and Usdin, T. (1994). Functional identification of a vesicular acetylcholine transporter and its expression from a "cholinergic" gene locus. J. Biol. Chem. *269*, 21929-21932.

Fazeli, A., Dickinson, S. L., Hermiston, M. L., Tighe, R. V., Steen, R. G., Small, C. G., Stoeckli, E. T., Keino-Masu, K., Masu, M., Rayburn, H., Simons, J., Bronson, R. T., Gordon, J. I., Tessier-Lavigne, M., and Weinberg, R. A. (1997). Phenotype of mice lacking functional Deleted in colorectal cancer (Dcc) gene. Nature *386*, 796-804.

Finney, M. and Ruvkun, G. (1990). The *unc-86* gene product couples cell lineage and cell identity in *C. elegans*. Cell *63*, 895-905.

Finney, M., Ruvkun, G., and Horvitz, H. R. (1988). The *C. elegans* cell lineage and differentiation gene *unc-86* encodes a protein with a homeodomain and extended similarity to transcription factors. Cell *55*, 757-769.

Fire, A. (1986). Integrative transformation of *Caenorhabditis elegans*. EMBO J. *5*, 2673-2680.

Garriga, G. and Stern, M. J. (1994). Hams and Egls: Genetic analysis of cell migration in *Caenorhabditis elegans*. Curr. Opin. Genet. Dev. *4*, 575-580.

Goshima, Y., Nakamura, F., Strittmatter, P., and Strittmatter, S. M. (1995). Collapsin-induced growth cone collapse mediated by an intracellular protein related to UNC-33. Nature *376*, 509-514.

Guo, S. and Kemphues, K. J. (1995). *par-1*, a Gene Required for Establishing Polarity in C. elegans Embryos, Encodes a Putative Ser/Thr Kinase That Is Asymmetrically Distributed. Cell *81*, 611-620.

Hall, D. H. and Hedgecock, E. M. (1991). Kinesin-related gene *unc-104* is required for axonal transport of synaptic vesicles in *C. elegans*. Cell *65*: 837-847.

Hamelin, M., Zhou, Y., Su, M. W., Scott, I. M., and Culotti, J. G. (1993). Expression of the UNC-5 guidance receptor in the touch neurons of *C. elegans* steers their axons dorsally. Nature *364*, 327-330.

Hata, Y., Slaughter, C. A., and Sühdof, T. C. (1993). Synaptic vesicle fusion complex contains unc-18 homologue bound to syntaxin. Nature *366*, 347-351.

Hedgecock, E. M., Culotti, J. G., and Hall, D. H. (1990). The unc-5, unc-6 and unc-40 genes guide circumferential migrations of pioneer axons and mesodermal cells on the epidermis in *C. elegans*. Neuron *2*, 61-85.

Hedgecock, E. M., Culotti, J. G., Thomson, J. N., and Perkins, L. A. (1985). Axonal guidance mutants of *Caenorhabditis elegans* identified by filling sensory neurons with fluorescein dyes. Dev. Biol. *111*, 158-170.

Hedgecock, E. M. and Hall, D. (1990). Homologies in the neurogenesis of nematodes, arthropods and chordates. Sem. Neurosci. *2*, 159-172.

Hedrick, L., Cho, K. R., Featon, E. R., Wu, T. C, Kinzler, K. W., and Vogelstein, B. (1994). The DCC gene product in cellular differentiation and colorectal tumorigenesis. Genes Dev. *8*, 1174-1183.

Heim, R., Cubitt, A. B., and Tsien, R. Y. (1995). Improved green fluorescence. Nature *373*, 663-664.

Higashide, T., Murakami, A., McLaren, M. J., and Inana, G. (1996). Cloning of the cDNA for a Novel Photoreceptor Protein. J. Biol. Chem. *271*, 1797-1804.

Hirai, H., Maru, Y., Hagiwara, K., Nishida, J., and Takaku, F. (1987). A novel putative tyrosine kinase receptor encoded by the *eph* gene. Science *238*, 1717-1720.

Hodgkin, J. (1980). More sex-determination mutants of *Caenorhabditis elegans*. Genetics *96*, 649-664.

Hodgkin, A. L. and Huxley, H. F. (1952). Currents carried by sodium and potassium ion through the membrane of the giant axon of *Loligo*. J. Physiol. *116*, 449-472.

Hodgkin, J. (1997). Appendix 1: Genetics. in *C. elegans II*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Hosono, R., Sassa, T., and Kuno, S. (1987). Mutations affcting acetylcholine levels in the nematode *Caenorhabditis elegans*. J. Neurochem. *49*, 1820-1823.

Ishii, N., Wadsworth, W. G., Stern, B. D., Culotti, J. G., and Hedgecock, E. M. (1992). UNC-6, a laminin-related protein, guides cell and pioneer axon migrations in *C. elegans*. Neuron *9*, 873-881.

Jansen, G., Hazendonk, E. Thijssen, K. L., and Plasterk R. H. A. (1997). Reverse genetics by chemical mutagenesis in *Caenorhabditis elegans*. Nature Genetics *17*, 119-121.

Jin, Y., Hoskins, R. and Horvitz, H. R. (1994). Control of type-D GABA ergic neuron differentiation by *C. elegans* UNC-30 homeodomain protein. Nature *372*, 780-783.

Johnson, C. D., Rand, J. B., Herman, R. K., Stern, B. D., and Russell, R. L. (1988). The acetylcholinesterase genes of *C. elegans*: Identification of a third gene (*ace-3*) and mosaic mapping of a synthetic lethal phenotype. Neuron *1*, 165-173.

Jorgensen, E. M., Hartwieg, E., Schuske, K., Nonet, M. L., Jin, Y. and Horvitz, H. R. (1995). Defective recycling of synaptic vesicles in synaptotagmin mutants of *Caenorhabditis elegans*. Nature *378*, 196-199.

Jorgensen, E. M. and Rankin, C. (1997). Neural Plasticity. in *C. elegans II*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Kaech, S., Ludin, B., and Matus, A. (1996). Cytoskeletal Plasticity in Cells Expressing Neuronal Microtubule-Associated Proteins. Neuron *17*, 1189-1199.

Kandel, E. R. (1976). *Cellular basis of behavior.* W. H. Freeman, San Francisco, CA.

Keino-Masu, K., Masu, M., Hinck, L., Leonardo, E. D., Chan, S. S. -Y., Culotti, J. G., and Tessier-Lavigne, M. (1996). Deleted in Colorectal Cancer (DCC) encodes a netrin receptor. Cell *87*, 175-185.

Kemphues, K. J., Priess, J. R., Morton, D. G., and Cheng, N. (1988). Identification of genes required for cytoplasmic localization in early *Caenorhabditis elegans* embryos. Cell *52*, 311-320.

Kennedy, T. E., Serafini, T., de la Torre, J. R., and Tessier-Lavigne, M. (1994). Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord. Cell *78*, 425-435.

Kenyon, C. (1997). Life Span. in *C. elegans II.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Keynes, R. and Cook, G. M. W. (1995). Axon Guidance Molecules. Cell *83*, 161-169.

Kirby, C., Kusch, M., and Kemphues, K. (1990). Mutations in the *par* genes of *Caenorhabditis elegans* affect cytoplasmic reorganization during the first cell cycle. Dev. Biol. *142*, 203-215.

Kristan, W. B. (1983). The neurobiology of swimming in the leech. Trends Neurosci. *6*, 84-88.

Kolodziej, P. A., Timpe, L. C., Mitchell, K. J., Fried, S. R., Goodman, C. S., Jan, L. Y., and Jan, Y. N. (1996). frazzled encodes a Drosophila member of the DCC immunoglobulin subfamily and is required for CNS and motor axon guidance. Cell *87*, 197-204.

Leonardo, E. D., Hinck, L., Masu, M., Keino-Masu, K., Ackerman, S. L., and Tessier-Lavigne, M. (1997). Vertebrate homologues of C. elegans UNC-5 are candidate netrin receptors. Nature *386*, 833-838.

Leung-Hagesteijn, C., Spence, A. M., Stern, B. D., Zhou, Y., Su, M. -W., Hedgecock, E. M., and Culotti, J. G. (1992). UNC-5, a transmembrane protein with immunoglobulin and thrombospondin type I domains, guides cell and pioneer axon migrations in *C. elegans.* Cell *71*, 289-299.

Lewis, J. A., Elmer, J. S., Skimming, J., McLafferty, S., Fleming, J., and McGee, T. (1987). Cholinergic receptor mutants of the nematode *Caenorhabditis elegans.* J. Neurosci. *7*, 3059-3071.

Lewis, J. A., Wu, C. H., Berg, H., and Levine, J. H. (1980). The genetics of levamisole resistance in the nematode *Caenorhabditis elegans*. Genetics *95*, 905-928.

Liu, K. S. and P. W. Sternberg. (1995). Sensory regulation of male mating behavior in *Caenorhabditis elegans*. Neuron *14*, 79-89.

Lockery, S. R. and Hall, D. H. (1995). Signal propagation in the nerve ring of the nematode *C. elegans*. Soc Neurosci. *21*, 569.7, p., 1454 (abstr.)

MacLeod, A. R., Karn, J., and Brenner, S. (1981). Molecular analysis of the *unc-54* myosin heavy chain gene of *Caenorhabditis elegans*. Nature *291*, 386-390.

Malone, E. A., Inoue, T., and Thomas, J. H. (1996). Genetic Analysis of the Roles of *daf-28* and *age-1* in Regulating *Caenorhabditis elegans* Dauer Formation. Genetics *143*, 1193-1205.

McIntire, S. L., Garriga, G., White, J., Jacobson, D., and Horvitz, H. R. (1992). Genes necessary for directed axonal elongation or fasciculation in *C. elegans*. Neuron *8*, 307-322.

McIntire, S. L., Jorgensen, E., Kaplan, J., and Horvitz, H. R. (1993). The GABAergic nervous system of *Caenorhabditis elegans*. Nature *364*, 337-341.

Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1992). Efficient gene transfer in *C. elegans*: Extrachromosomal maintenance and integration of transforming sequences. EMBO J. *10*, 3959-3970.

Miller, D .M. and Niemeyer, C. J. (1995). Expression of the *unc-4* homeoprotein in *Caenorhabditis elegans* motor proteins specifies presynaptic input. Development *121*, 2877-2886.

Miller, D. M., Shen, M. M., Shamu, C. E., Bürglin, T. R., Ruvkun, G., Dubois, M. L., Ghee, M., and Wilson, L. (1992). *C. elegans unc-4* gene encodes a homeodomain protein that determines the pattern of synaptic input to specific motor neurons. Nature *355*, 841-845.

Monschau, B., Kremoser C., Ohta K., Tanaka H., Kaneko T., Yamada T., Handwerker C., Hornberger, M. R., Löschinger J., Pasquale E. B., Siever D. A., Verderame M. F., Müller B. K., Bonhoeffer F., and Drescher, U. (1997). Shared and distinct functions of RAGS and ELF-1 in guiding retinal axons. EMBO J *16(6)*,1258-1267.

Moskowitz, I. P., Gendreau, S. B., and Rothman, J. H. (1994). Combinatorial specification of blastomere identity by *glp-1*-dependent cellular interactions in the nematode *Caenorhabditis elegans*. Development *120*, 3325-3338.

Müller, B. K., Bonhoeffer, F., and Drescher, U. (1996). Novel gene families involved in neural pathfinding. Curr. Opinion Gen. Dev. *6*, 469-474.

Nakamoto, M., Cheng, H.-J., Friedman, G. C., McLaughlin, T., Hansen, M. J., Yoon, C. H., O'Leary, D. D. M., and Flanagan, J. G. (1996). Topographically specific effects of ELF-1 on retinal axon guidance in vitro and retinal axon guidance in vivo. Cell *86*, 755-766.

Nguyen, M., Alfonso, A., Johnson, C. D., and Rand, J. B. (1995). *Caenorhabditis elegans* mutants resistant to inhibitors of acetylcholinesterase. Genetics *140*, 527-535.

Nicholls, J. G., Martin, A. R., and Wallace, B. G. (eds.) (1992). *From neuron to brain: a cellular and molecular approach to the function of the nervous system*, 3rd ed. Sinauer Associates, Inc. Sunderland, Massachusetts.

Nieto, M. A. (1996). Molecular Biology of Axon Guidance. Neuron *17*, 1039-1048.

Nomarski, G. (1955). Microinterféromètre differentiel á ondes polarisées. J. Phys. Radium *16*, 9-13.

Nonet, M. L., Grundahl, K., Meyer, B. J., and Rand, J. B. (1993). Synaptic function is impaired but not eliminated in *C. elegans* mutants lacking synaptotagmin. Cell *73*, 1291-1305.

Ogura, K., Wicky, C., Magnenat, L., Tobler, H., Mori, I., Muller, F., and Ohshima, Y. (1994). *Caenorhabditis elegans unc-51* gene required for axonal elongation encodes a novel serine/threonine kinase. Genes Dev. *84*, 2389-2400.

Ogura, K., Shirakawa, M., Barnes, T. M., Hekimi, S., and Ohshima, Y. (1997). The UNC-14 protein required for axonal elongation and guidance in *Caenorhabditis elegans* interacts with the serine/threonine kinase UNC-51. Genes Dev. *11*, 1801-1811.

Otsuka, A. J., Franco, R., Yang, B., Shim, K. -H., Tang, L. Z., Zhang, Y. Y., Boontrakulpoontawee, P., Jeyaprakash, A., Hedgecock, E., Wheaton, V. I., and Sobery, A. (1995). An Ankyrin-related Gene (*unc-44*) Is Necessary for Proper Axonal Guidance in *Caenorhabditis elegans*. J. Cell Biol. *129*, 1081-1092.

Otsuka, A. J., Jeyaprakash, A., Garcia-Anoveros, J., Tang, L. Z., Fisk, G., Hartshorne, T., Franco, R., and Born, T. (1991). The *C. elegans unc-104* gene encodes a putative knesin heavy chain-like protein. Neuron *6*, 113-122.

Perkins, L. A., Hedgecock, E. M., Thomson, J. N., and Culotti, J. G. (1986). Mutant sensory cilia in the nematode *Caenorhabditis elegans.* Dev. Biol. *117*, 456-487.

Plasterk, R. H. A. (1995). Reverse Genetics: From Gene Sequence to Mutant Worm. in *Caenorhabditis elegans*: *Modern Biological Analysis of an Organism* (eds Epstein, H. F. and Shakes, D. C.) Methods Cell Biol. *48*, 59-80.

Plasterk, R. H. A. and Groenen, J. T. M. (1992). Targeted alterations of the *Caenorhabditis elegans* genome by transgene instructed DNA double strand break repair following Tc1 excision. EMBO J. *11*, 287-290.

Ramón y Cajal, S. (1890). Sobre la aparición de las expansiones celulares en la médula embrionaria. Gaceta Sanitaria de Barcelona *12*, 413-419.

Ramón y Cajal, S. (1892). La rétine des vertébrés. La cellule, *IX*, 119-258.

Rand, J. B. (1989). Genetic analysis of the *cha-1-unc-17* gene complex in *Caenorhabditis.* Genetics *122*, 73-80.

Rand, J. B. and Nonet, M. L. (1997a). Appendix 2: Neurotransmitter Assignment for Specific Neurons. in *C. elegans II.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Rand, J. B. and Nonet, M. L. (1997b). Synaptic Transmission. in *C. elegans II.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Ridley, A. J., and Hall, A. (1992). The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors. Cell *70*, 389-399.

Ruvkun, G. (1997). Patterning the Nervous System. in *C. elegans II.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Serafini, T., Kennedy, T. E., Galko, M. J., Mirzayan, C., Jessell, T. M., and Tessier-Lavigne, M. (1994). The netrins define a family of axon outgrowth-promoting proteins homologous to *C. elegans* UNC-6. Cell *78*, 409-424.

Starich, T. A., Herman, R. K., Kari, C. K., Yeh, W. -H., Schackwitz, W. S., Schuyler, M. W., Collet, J., Thomas, J. H., and Riddle, D. L. (1995). Mutations affecting the chemosensory neurons of *Caenorhabditis elegans.* Genetics *139*, 171-188.

Stoeckli, E. T., Sonderegger, P., Pollerberg, G. E., and Landmesser, L. T. (1997). Interference with axonin-1 and NrCAM interactions unmasks a floor-plate activity inhibitory for commissural axons. Neuron *18*, 209-221.

Stossel, T. P. (1993). On the Crawling of Animal Cells. Science *260*, 1086-1094.

Stretton, A. O. W., Fishpool, R. M., Southgate, E., Donomoyer, J. E., Walrond, J. P., Moses, J. E., and Kass, I. S. (1978). Structure and physiological activity of the motorneurons of the nematode *Ascaris*. PNAS(USA) *75*, 3493-3497.

Südhof, T. C. (1995). The synaptic vesicle cycle: a cascade of protein-protein interactions. Nature *375*, 645-653.

Sulston, J .E. (1976). Post-embryonic development in the ventral nerve cord of *Caenorhabditis elegans*. Philos. Trans. R. Soc. Lond. B Biol. Sci. *275*, 287-298.

Sulston, J. E. (1988). Cell Lineage. in *The Nematode Caenorhabditis elegans*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Sulston, J. E. and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. Dev. Biol. *56*, 110-156.

Sulston, J. E. and Horvitz, H. R. (1981). Abnormal cell lineages in mutants of the nematode *Caenorhabditis elegans*. Dev. Biol. *82*, 41-55.

Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. Dev. Biol. *100*, 64-119.

Sulston, J., Du, Z., Thomas, K., Wilson, R., Hillier, L., Staden, R., Halloran, N., Green, P., Thierry-Mieg, J., Qiu, L., Dear, S., Coulson, A., Craxton, M., Durbin, R., Berks, M., Metstein, M., Hawkins, T., Ainscough, R., and Waterston, R. (1992). The *C. elegans* genome sequencing project: A begininning. Nature *356*, 37-41.

Tanaka, E. and Sabry, J. (1995). Making the Connection: Cytoskeletal Rearrangements during Growth Cone Guidance. Cell *83*, 171-176.

Taylor, P. and Radi, Z. (1994). The cholinesterases: From genes to proteins. Annu. Rev. Pharmacol. Toxicol. *34*, 281-320.

Tessier-Lavigne, M., Placzek, M., Lumsden, A. G. S., Dodd, J., and Jessell, T. M. (1988). Chemotropic guidance of developing axons in the mammalian central nervous system. Nature *336*, 775-778.

Tessier-Lavigne, M. and Goodman, C. S. (1996). The Molecular Biology of Axon Guidance. Science *274*, 1123-1133.

Thomas, J. H. (1990). Genetics analysis of defecation in *Caenorhabditis elegans*. Genetics *124*, 855-872.

Timpl, R. and Brown, J. C. (1996). Supramolecular assembly of basement membranes. BioEssays *18*, 123-132.

Treinin, M. and Chalfie, M. (1995). A mutated acetylcholine receptor subunit causes neuronal degeneration in *C. elegans.* Neuron *14*, 871-877.

Trent, C., Tsung, N., and Horvitz, H. R. (1983). Egg-laying defective mutants of the nematode *Caenorhabditis elegans.* Genetics *104*, 619-647.

Tully, T., Preat, T., Boynton, S. C., Del Vecchio, M. (1994). Genetic dissection of consolidated memory in Drosophila. Cell *79*, 35-47.

Wadsworth, W. G., Bhatt, H., and Hedgecock, E. M. (1996). Neuroglia and Pioneer Neurons Express UNC-6 to Provide Global and Local Netrin Cues for Guiding Migrations in C. elegans. Neuron *16*, 35-46.

Wadsworth, W. G. and E. M. Hedgecock. (1996). Hierarchical guidance cues in the developing nervous system of *C. elegans.* BioEssays *18(5)*, 355-362.

Walter, J., Kern-Veits, B., Huf, J., Stolze, B., and Bonhoeffer, F. (1987). Recognition of position-specif properties of tectal membranes by retinal axons in vitro. Development *101*, 685-696.

Waterston, R. H., Sulston, J. E., and Coulson, A. R. (1997). The Genome. in *C. elegans II.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Waterston, R. H. (1988). Muscle. in *The nematode* Caenorhabditis elegans. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Weinshenker, D., G. Garriga, and Thomas, J. H. (1995). Genetic and Pharmacological Analysis of Neurotransmitters Controlling Egg Laying in *C. elegans.* J. Neuroscience *15(10)*, 6975-6985.

White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1976). The structure of the ventral nerve cord of *Caenorhabditis elegans.* Phil. Trans. R. Soc. Lond. B Biol. Sci. *275*, 327-348.

White, J. G., Albertson, D. G., and Anness, M. A. R. (1978). Connectivity changes in a class of motoneurone during the development of a nematode. Nature *271*, 764-766.

White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1986). The structure of the nervous system of the nematode *C. elegans.* Phil. Trans. R. Soc. Lond. B Biol. Sci. *314*, 1-340.

White, J., Southgate, E., and Durbin, R. (1988). Appendix 2: Neuroanatomy. in *The Nematode Caenorhabditis elegans.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

White, J., Southgate, E., and Thomson, J. N. (1992). Mutations in the *Caenorhabditis elegans unc-4* gene alter the syanptic input to ventral cord motor neurons. Nature *355*, 838-841.

Wicks, S. R. and Rankin, C. H. (1995). Integration of mechanosensory stimuli in *Caenorhabditis elegans.* J. Neurosci. *15*, 2434-2444.

Wicks, S. R. and Rankin, C. H. (1996). The integration of antagonistic reflexes revealed by laser ablation of identified neurons determines habituation kinetics of the *Caenorhabditis elegans* tap withdrawal response. J. Comp. Physiol. A *179*, 675.

Wightman, B., Baran, R., and Garriga, G. (1997). Genes that guide growth cones along the *C. elegans* ventral nerve cord. Development *124*, 2571-2580.

Wightman, B., Clark, S. G., Taskar, A. M., Forrester, W. C., Maricq, A. V., Bargmann, C. I., and Garriga, G. (1996). The *C. elegans* gene *vab-8* guides posteriorly directed axon outgrowth and cell migration. Development *122*, 671-682.

Wood, W. B. (ed.) (1988). *The Nematode Caenorhabditis elegans.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Yandell, M. D., Edgar, L. G., and Wood, W. B. (1994). Trimethylpsoralen induces small deletion mutations in *Caenorhabditis elegans.* PNAS(USA) *91*, 1381-1385.

Zarkower, D., and Hodgkin, J. (1992). Molecular analysis of the *C. elegans* sex-determining gene *tra-1*: A gene encoding two zinc finger proteins. Cell *70*, 237-249.

Zipkin, I., Kindt, R. M., and Kenyon, C. J. (1997). Role of a New Rho Family Member in Cell Migration and Axon Guidance in *C. elegans*. Cell *90*, 883-894.

2 Identification and Cloning of *unc-119*, a Gene Expressed in the *Caenorhabditis elegans* Nervous System

2.1 Introduction

As foreseen by Brenner over twenty years ago (Brenner, 1974), the free-living nematode *Caenorhabditis elegans* is an ideal system in which to study the development of a complete nervous system. Development in *C. elegans* follows a nearly invariant lineage, and the locations and patterns of connectivity of neurons are similar in every animal. Of the approximately 1000 somatic nuclei in the adult hermaphrodite, 37% comprise the nervous system (Chalfie and White, 1988), including 302 neurons. Serial-section electron micrograph reconstructions have allowed the exact placement of each neuron cell body and axonal projections (White *et al.*, 1986). *C. elegans* has two independent nervous systems: Feeding is controlled by the pharyngeal nervous system (Albertson and Thomson, 1976), while movement and behavior are controlled by the somatic nervous system (White *et al.*, 1986). Since *C. elegans* is capable of self-fertilization, animals do not need to move in order to reproduce. These properties, coupled with the ease of *C. elegans* genetic manipulation (Brenner, 1974), have led to the identification of over one hundred genes involved in the development of the muscle and nervous systems (Wood, 1988).

Over thirty different genes required for normal muscle position, structure or function have been identified (Anderson, 1989), many of which confer an

---

A version of this chapter has been published. Maduro M and Pilgrim D (1995). Genetics 141: 977-988.

"Unc" mutant phenotype (for uncoordinated), such as myosin heavy chain (Macleod *et al.*, 1989). Mutations in genes affecting the nervous system often have a more subtle Unc phenotype, affecting only a small subset of behaviors, such as response to gentle touch, sensitivity to volatile attractants, contractile ability of certain sets of body muscles and constitutivity of (normally) regulated behavior such as feeding or defecation. To date, several of these have been cloned and characterized, including kinesin homologues (Hall and Hedgecock, 1991), clathrin-associated proteins (Lee *et al.*, 1994), and cell adhesion and membrane proteins involved in nerve axon guidance and pathfinding (Hedgecock *et al.*, 1990; Ishii *et al.*, 1992; Leung-Hagesteijn *et al.*, 1992; Rogalski *et al.*, 1993). In some cases, human homologues have recently been identified for gene products previously known only in the *C. elegans* nervous system (Hosono *et al.*, 1992; Hata *et al.*, 1993).

In this paper, we report the identification of a new gene, *unc-119*, whose defect is likely neuronal based on both indirect and direct evidence. Animals homozygous for putative null alleles display many abnormalities in neuron-mediated behaviors; examination of muscle shows normal ultrastructure; and expression of a reporter gene fused to *unc-119* shows expression in different types of neurons. It is likely, therefore, that the *unc-119* gene product is required for the establishment or function of the nervous system. The predicted UNC-119 protein shows no significant similarity to other known proteins except for another *C. elegans* gene, of unknown function, suggesting that there may be a family of similar proteins involved in the nervous system.

2.2 Materials and Methods

2.2.1 *Strains and Genetics*

Growth, maintenance and manipulation of strains was performed as described (Wood, 1988). All strains were obtained from the stock collection of the MRC Laboratory of Molecular Biology, Cambridge, UK (with the help of J. Hodgkin), with the exception of: *dpy-28(y1) III* and *daf-11(m47ts) V* which were obtained from the Caenorhabditis Genetics Center; RW7097 [*mut-6(st702) unc-22(st192st527) IV*] which was a gift from B. Williams; and DA438 [*bli-4(e937) I*; *rol-6(e187) II*; *daf-2(e1368) vab-7(e1562) III*; *unc-31(e928) IV*; *dpy-11(e224) V*; *lon-2(e678) X*] which was a gift from L. Avery (Avery, 1993).

2.2.2 *Generation of unc-119 alleles*

The initial *unc-119* mutation, *e2498*, arose spontaneously in the progeny of a hybrid between two *C. elegans* strains, Bristol (N2) and Bergerac (BO) (Pilgrim, 1993). *unc-119(ed3)* and *unc-119(ed4)* were isolated in the F_1 generation of 1500 *tDf2 + / + dpy-18(e499)* hermaphrodites mutagenized in 100 mM EMS (Sigma). *unc-119(ed9)* was isolated in a similar screen. The revertant DP41 [*unc-119(e2498ed5)*] arose spontaneously from a strain carrying both *e2498* and *mut-6(st702)*.

2.2.3 *Mapping of unc-119*

The gene *unc-119* was initially mapped to chromosome *III* using the mapping strain DA438. Crosses with *daf-7, unc-64, dpy-1* and *lon-1* placed *unc-119* on the right half of chromosome *III*, between *lon-1* and *unc-64*. *unc-119* complements other Unc genes in this region including *unc-25*, *unc-32*, *unc-47*, *unc-49*, *unc-71* and *unc-81*.

a) *unc-119* maps very close to *vab-7*. Dpy non-Vab and Vab non-Dpy self-progeny were selected from a hermaphrodite of genotype *vab-7(e1562) dpy-18(e499)* + / + + *unc-119*. These recombinant progeny fell into the following classes: *vab-7 dpy-18* / + *dpy-18* (0 animals); *vab-7 dpy-18* + / + *dpy-18 unc-119* (31 animals); *vab-7 dpy-18* / *vab-7* + (35 animals); and *vab-7 dpy-18* + / *vab-7* + *unc-119* (0 animals). Sma non-Vab self-progeny were selected from a hermaphrodite of genotype *sma-2(e502) vab-7(e1562)* + / + + *unc-119(e2498)*. The progeny consisted of *sma-2 vab-7* + / *sma-2* + *unc-119* (14 animals) and *sma-2 vab-7* / *sma-2* + (0 animals).

b) *unc-119* is deleted by *tDf2*. *unc-119(e2498)* heterozygous males were crossed to hermaphrodites of genotype *tDf2* + / + *dpy-18(e499)*. Unc non-Dpy progeny were seen in the F_1 brood.

c) *unc-119* maps to the break between *eDf2* and *eDp6*, as neither rearrangement complements *unc-119*. *unc-119(e2498) dpy-18(e499)* heterozygous males were crossed to hermaphrodites of strain CB1517 [*eDf2; eDp6*]. Wild-type progeny were recovered in the F_1, but none carried the *unc-119 dpy-18* chromosome, as judged by progeny testing, suggesting that neither *eDf2* nor *eDp6* complement *unc-119*. Unc-non-Dpy animals were recovered among the F_1 outcross progeny, and shown to carry *eDf2; eDp6* and the *unc-119 dpy-18* chromosome by progeny testing. The chromosomal break that led to the production of *eDf2* and *eDp6* maps just to the left of *vab-7* (Hodgkin, 1980), which places *unc-119* at this position.

2.2.4 Construction of unc-119(e2498); daf-7(e1372ts) and unc-119(e2498); daf-11(m47ts) double mutants

Heterozygous *unc/+* males were crossed to adult *daf* hermaphrodites and placed at 25° after mating. F_1 animals were selfed at 25° and dauer progeny (scored by morphology) were picked from the F_2 brood to recover and self-fertilize at 15°. Unc animals were then picked from the F_3 brood. Verification of the double mutant genotype was made by crossing to wild type and checking the F_2 segregation proportions.

2.2.5 Phenotypic Analysis

Pharyngeal pumping, egg retention and locomotion were scored as in Ségalat *et al*, 1995, with modifications. Locomotion was scored as complete sinusoidal waves per minute in liquid by adding 20 µL M9 buffer to a plate of worms and observing the motion under a dissecting microscope. To score egg retention, well-fed adult hermaphrodites were observed using Differential Interference Contrast (DIC) optics. Pharyngeal pumping was scored by observing contractions of the terminal bulb of the pharynx for one minute and then transferring the same animals to an unseeded plate. The animals were rescored after five minutes and then returned to a seeded plate to verify that pumping resumed to 'on food' levels. For the transgenic strains *unc-119(e2498); edEx31* [pDP#MM016], *unc-119(ed3)*; *edEx32* [pDP#MM051], and *unc-119(ed4)*; *edEx33* [pDP#MM019], the array is extrachromosomal, with some meiotic and mitotic loss, and animals with the best visible locomotion were chosen for scoring of all traits, to assure at least partial inheritance of the transgenic array. At least ten animals were scored for each trait.

Ability to form dauer larvae was initially measured by determining resistance to SDS in starved cultures at 25° (Cassada and Russell, 1975). Plates on which animals had depleted the supply of bacteria were left for four days. At least five thousand worms per plate were collected, incubated in 2% SDS (in M9) for 2 hours at 25°, rinsed and dropped onto an unseeded plate. The ability to form dauers was scored as positive if surviving animals were seen. In all such cases, there were at least fifty survivors per original plate, and they had an appearance and behavior characteristic of dauer larvae. When fed bacteria, most of these animals recovered to form adults (at 15° in the case of *unc-119; daf-7*). Plates of strains that were unable to form SDS-resistant larvae also lacked animals with characteristic dauer morphology.

2.2.6 Molecular Analysis

Genomic DNA was prepared as described (Pilgrim, 1993). The λZAP cDNA library was a gift from B. Barstead (Barstead and Waterston, 1989). Screening of the library and *in vivo* excision was performed as described by the manufacturer (Stratagene). Cosmids and YACs from the physical map of *C. elegans* were obtained from A. Coulson. A clone containing the *C. elegans* transposable element Tc1 was a gift from L. Harris. The yeast strain N123 was obtained from R. C. von Borstel.

Gel electrophoresis and Southern Blotting was performed as described (Pilgrim, 1993). All cloning, unless otherwise stated, was performed in pBluescript II KS- (Stratagene) grown in *E. coli* XL1-B. Manipulation of DNA was performed as described (Sambrook *et al*, 1989). All sequencing was done on single-stranded templates as described in the USB Sequenase 2.0 kit (United States Biochemicals).

The *lacZ* fusion construct was made using the vector pPD22.04 (Fire *et al*, 1990). A 3.6 kbp *Hin*d III - *Pst* I fragment was purified from pDP#MM016 (see Figure 2.7) and ligated to similarly digested pPD22.04 to produce the *unc-119::lacZ* fusion plasmid. The fusion is predicted to contain the first 101 amino acids of UNC-119.

PCR was performed using the primers MMA1 (5'-AGTCGGCCTTATTGT GCATTAC-3') and MMA2 (5'-AAATTGCATGCCAGCACCGGTC-3'), on ~100 ng of genomic DNA of alleles *ed3, ed4, ed9* and *e2498ed5* in a 25μl volume on a Stratagene Robocycler 40 using the following conditions: one cycle (95°, 60 sec; 61°, 90 sec; 73°, 180 sec) and 29 cycles (93°, 75 sec; 61°, 75 sec; 73°, 135 sec). PCR products were purified, digested with *Nhe* I and *Sst* I, and subcloned into pBluescript II SK- and KS- (Stratagene) that had been digested with *Xba* I and *Sst* I (all enzymes BRL). Single strands were packaged prior to sequencing. Two independent PCR products were subcloned and sequenced in each case.

RACE (Rapid Amplification of cDNA Ends; Frohman *et al.*, 1988) was performed using a 5' RACE System kit (BRL) according to the manufacturer's instructions. The primer used for first strand synthesis was MMA1 (sequence shown above) and the nested primer used was MMA5 (5'- GGAGCATAGGAA TTCTTGAGTGATTCC-3'; the underlined base indicates a single mismatch that introduces an *Eco* RI site). The initial RACE product was reamplified using DHA16 (5'-TACTAGGCCACGCGTCGACTAGTA-3'; this primer contains part of the sequence found in the UAP primer supplied with the kit) and MMA5. The single 300 bp product was purified, digested with *Eco* RI (BRL) and *Sal* I (Pharmacia), and cloned into similarly digested pBluescript II SK- for sequencing. Multiple clones of two independent products were sequenced.

Worm and yeast plugs were prepared and handled as described (Birren and Lai, 1993). Restriction digestions were performed using *Not* I (BRL) after treatment of plugs with 100 mM PMSF (Sigma). Pulsed-Field Gels were electrophoresed in 1% agarose (BRL) in a Hex-A-Field CHEF gel apparatus (BRL) for 16 hours at 150 V with a switching time of 60 sec. Gels were run in 0.5 x TBE running buffer cooled by circulation through a glass coil immersed in a cooling bath set to 8°. YAC DNA was purified by running 1% low gel temperature agarose gels (BRL) and cutting out the correct band after staining in 200 mL of 0.5 µg/mL Ethidium Bromide in 0.5 x TBE for 30 minutes at room temperature and visualizing bands under 254-nm light using a transilluminator. Agarose was removed using GELase (Cedarlane Laboratories) as per the instructions accompanying the enzyme.

2.2.7 Microscopy and C. elegans transformation

Generation of transgenic animals by microinjection of DNA into worm gonads was done as described (Mello *et al.*, 1991), using a Zeiss Axiovert microscope equipped with DIC optics. The plasmid pRF4, which contains the dominant mutation *rol-6(su1006)*, was used as a transformation control. Animals transgenic for the *unc-119::lacZ* fusion were fixed and stained for β-galactosidase activity as described (Fire *et al.*, 1990).

Electron microscopy was performed on a Philips EM201 transmission electron microscope. Embedding was performed with Epon after fixation with glutaraldehyde and osmium tetroxide as described (Wood, 1988). 100-nm sections were cut using a Reichert OMU2 ultramicrotome using glass knives prepared on an LKB Knifemaker II. Sections were collected on 300-mesh copper grids.

2.3 Results

2.3.1 Identification and mutant phenotype of unc-119

The *e2498* mutation arose spontaneously as a strong uncoordinated animal among the progeny of a cross between two nematode strains, Bristol (N2) and Bergerac (BO) during a search for strain polymorphisms on chromosome *III* (Pilgrim, 1993). Interstrain crosses of this type are known to increase mobility of the transposable element Tc1 (Mori *et al.*, 1990). The Unc phenotype was completely recessive, visible from hatching and remained strong after backcrossing ten times to N2. Older *e2498* animals are slightly dumpier in shape than wild type, and while they can forage with their heads and appear to respond to touch, mutants usually tend to remain in one position on the plate, often forming a coil or limp 'C' shape (Figure 2.1). When locomotion is assayed by scoring sinusoidal waves per minute in liquid (see Materials and Methods), mutants cannot complete more than ten body bends per minute, as compared to wild type, which can complete more than 100 (Figure 2.2). When exposed to levamisole (an acetylcholine agonist), both wild type animals and homozygous *e2498* hypercontract, suggesting that acetylcholine receptors in *e2498* are normal (Lewis *et al.*, 1980). This also suggests that the defect in *unc-119* mutants lies in the nervous system rather than the muscle system. Males homozygous for *e2498* are similarly paralyzed, but show normal bursa morphology and spicule movement under DIC optics. Hermaphrodites are still capable of self-fertilization and give reduced brood sizes about 50% of wild type, and the development time from embryo to adult is not affected.

Further evidence as to a role in the nervous system comes from analysis of neuron-mediated traits as shown in Figure 2.2. Under overcrowded conditions, or in the presence of dauer pheromone, wild type nematodes can enter the dauer state, a developmentally arrested and morphologically distinct alternative third larval stage (Cassada and Russell, 1975). *unc-119* mutants cannot form such larvae as assayed by testing starved plates for SDS-resistance (see Materials and Methods). We tested for the persistence of this dauer-forming defect (Daf-d phenotype) in combination with *daf-7(e1372ts)* and *daf-11(m47ts)*, two mutations that cause constitutive dauer formation at 25° (a Daf-c phenotype; Riddle *et al.*, 1981). *unc-119* is able to suppress the Daf-c phenotype of *daf-11* as *unc-119(e2498); daf-11(m47ts)* animals are Daf-d at 25°. However, *unc-119(e2498); daf-7(e1372ts)* animals are Daf-c at 25°, suggesting that the Daf-d phenotype of *unc-119* mutants does not result from a physiological inability to enter dauer.

Other behaviors are similarly affected. When shifted from a bacterial lawn to an unseeded plate, wild-type animals will dramatically slow the rate of pharyngeal pumping (Avery and Horvitz, 1990). In contrast, while the rate of pumping is similar to wild type in the presence of food, *e2498* animals continue to pump their pharynxes off food at a moderately high rate, albeit somewhat reduced. Wild-type animals also cease egg laying in the absence of food (Avery and Horvitz, 1990) and on average retain a consistent number of eggs (Ségalat *et al.*, 1995). *unc-119* mutants tend to retain more embryos than wild type, and when they do lay eggs, they can do so in the absence of food, likely owing to increased pressure inside the uterus. Constitutive pharyngeal pumping, an egg laying defect and an apparent inability to enter dauer all suggest that *unc-119* has a role in the nervous system.

To rule out defects in musculature, *unc-119(2498)* adults were fixed and sectioned for transmission electron microscope analysis of body wall muscle structure. Muscle ultrastructure, placement of sarcomeres and attachment to the nematode body wall appear normal. Mutant animals were also examined with polarized light, which can reveal defects in musculature (Epstein and Thomson, 1974), as well as with indirect immunofluorescence with an anti-vinculin antibody, which stains the thin filaments. In all muscle characteristics examined, no differences were seen compared to wild-type controls (data not shown).

2.3.2 Mapping of unc-119

The mutation *e2498* was mapped to chromosome *III* using the multiply marked mapping strain DA438 and positioned with respect to several known markers in the area (see Materials and Methods). The new mutation is between *sma-2* and *dpy-18*, and is almost inseparable from *vab-7*. *e2498* is not complemented by either deficiency *eDf2* or *tDf2* (Schnabel and Schnabel, 1990) and is not contained in the free duplication *eDp6*, although all other known genes in the area are complemented by either the *eDf2* chromosome or the *eDp6* duplication. This places *unc-119* just to the left of *vab-7*, and to the right of *dpy-28* (Figure 2.3). Since there were no known *unc* genes at this position, the mutation was assigned to a new gene, *unc-119*. J. Hodgkin noticed that the phenotype of *unc-119* homozygotes was similar to that of strain CB1517, which is homozygous for *eDf2* and *eDp6*. This suggests that *unc-119* is at least partially deleted in CB1517. *eDf2* and *eDp6* are chromosome *III* derived aberrations obtained concomitantly after acetaldehyde mutagenesis (Hodgkin, 1980). Presumably the mutagenesis created a chromosomal break at the site, that healed to allow *eDf2* and *eDp6* to segregate independently from one another. Since *C. elegans* chromosomes are mitotically holocentric (Albertson and Thomson, 1982), both products of such a rearrangement can be maintained.

eDf2 and *eDp6* can be seen as bivalents using fluorescent *in situ* hybridization (Albertson, 1993). CB1517 animals are viable and otherwise healthy, although they do show reduced brood sizes due to meiotic loss of *eDp6*. Since the phenotypes of *e2498* and *eDf2; eDp6* are almost indistinguishable (see Figures 1 and 2), *unc-119* may be the only gene affected in these animals.

Three new alleles of *unc-119* (*ed3, ed4* and *ed9*) were recovered in the F_1 progeny of EMS-mutagenized animals heterozygous for *tDf2*. These new alleles arose at a forward mutation frequency of ~2 x 10^{-5}. This figure is approximately tenfold lower than the expected value for an average-sized gene in *C. elegans* (Brenner, 1974), suggesting that either the gene is small or that few alleles are viable or have a visible phenotype. The new alleles are completely recessive, fail to complement *e2498*, *eDf2* or *eDp6* and are phenotypically similar to *e2498* (see Figure 2.2). Homozygotes are indistinguishable from hemizygotes, when *unc-119* is placed opposite the deficiency *tDf2*, suggesting the alleles may be null.

2.3.3 Cloning of unc-119

Since *unc-119(e2498)* was derived from a dysgenic cross, it seemed possible that it was due to the insertion of a transposable element into a single gene. *mut-6(st702) IV*, a presumed transposase source that causes high rates of mobilization of the transposable element Tc1 (Mori *et al.*, 1990), was introduced into a strain carrying *e2498*, in an attempt to recover revertants due to transposon excision. Homozygous *unc-119(e2498); mut-6(st702)* animals were cultured for several generations, and wild-type animals were found on a single plate. These wild-type animals segregated wild-type and *unc-119* progeny, and the suppressor phenotype mapped to the *unc-119* gene (data not shown). The putative revertant, *unc-119(e2498ed5)*, when homozygous, is indistinguishable from wild type and does not resegregate the *unc-119* phenotype, suggesting that reversion was intragenic, and may have been due to precise excision of a transposon insertion. Although 'transposon tagging' has been used to isolate other Unc genes in *C. elegans* (Moerman *et al.*, 1986), we were unable to detect an allele-specific polymorphism on Southern blots of *unc-119(e2498)* genomic DNA probed with Tc1 or any other transposable element, due to the background of other Tc1-hybridizing bands (data not shown).

The genome of *C. elegans* has been fractionated into overlapping Yeast Artificial Chromosome (YAC) and cosmid clones, and a physical map of the genome has been assembled (Coulson *et al.*, 1986; 1988; 1991). From the physical map, four YACs and several cosmid clones were selected for further analysis by their proximity to the cloned loci *tra-1* and *pha-1* (Hodgkin, 1993; Granato *et al.*, 1994). Since it was assumed that *unc-119* is disrupted or absent from both the *eDf2* deficiency and the *eDp6* duplication, and the breakpoints of the two rearrangements lie just to the left of *vab-7*, mapping the breakpoints of the two rearrangements should define the sequences necessary for *unc-119* function. T. Barnes suggested to us that Southern blots prepared from genomic DNA of strains carrying varying numbers of the duplications and deficiencies would reveal, by dosage, whether or not a clone was to the left or right of *unc-119* (see Figure 2.4). As judged by dosage on such blots, cosmids K02B6 and T27G1 contain DNA to the left of *unc-119* and T19F2, W06D6, W09D6, M02H11 and W03F10 lie to the right. None of these cosmids apparently contain DNA from both *eDf2* and *eDp6*.

The YAC clones were used for long-range restriction mapping of the genomic region. All the YACs overlap with a 300 kbp *Not* I fragment in genomic DNA from the wild type N2 strain, as determined by cross-hybridization of the cosmid clones to *Not* I digested YAC and worm DNA (Figure 2.5). By correlating the size of the cross-hybridizing *Not* I fragments in the YACs with the size of the YACs themselves, these clones were positioned to within 10 kbp. The DNA from a strain homozygous for both *eDf2* and *eDp6* contains an *eDf2*-specific *Not* I fragment on Southern blots that is larger than the N2 fragment, demonstrating that extra DNA, perhaps telomeric, has been added beyond the breakpoint. The size of the *eDp6*-specific *Not* I fragment implies that if the left end of the fragment is the actual end of *eDp6* itself, very little DNA has been added.

Based on the restriction map evidence of the YAC clones tested, only Y60D9 contains DNA from both *eDf2* and *eDp6*. Fluorescent *in situ* hybridization (FISH) has shown that the YAC clone Y39A1 cross-hybridizes to the meiotically visible *eDf2* bivalent and the *eDp6* chromosome (Albertson, 1993), implying that this YAC clone must contain *unc-119* and suggesting that its right endpoint is in between the leftmost *eDp6*-hybridizing cosmids including T19F2 (since all of the *eDp6*-containing cosmids fail to cross-hybridize Y39A1) and the breakpoint itself. Plasmid subclone libraries were constructed from Y39A1 and Y60D9 DNA after enrichment for DNA from these YACs by pulsed-field gel purification. Clones were tested for worm specificity by cross-hybridization against "spot blots" of genomic DNA from the *S. cerevisiae* strain N123 (which contains no YACs), Y42D3 (a YAC from the left of chromosome *III*), the other YACs shown in Figure 2.5 and wild-type *C. elegans* genomic DNA. Of the first 20 insert-containing clones picked from the library, four were found to be worm-specific and contained in Y39A1 and Y60D9 alone (data not shown). One subclone, pDP#MM008, was used to probe Southern blots of genomic digests of N2, BO, *unc-119(e2498)*, CB1517 [*eDf2; eDp6*] and the revertant *unc-119(e2498ed5)* (see Figure 2.6). This clone detects a 4.5 kbp *Sst* I fragment in N2 and BO genomic DNA (these are the parent strains of the original *e2498* mutation) and a larger, 6.1 kbp fragment in *e2498*. The revertant allele *e2498ed5* contains the normal 4.5 kbp band, consistent with pDP#MM008 being an *unc-119* specific clone, and that *unc-119(e2498)* resulted from the insertion of a 1.6 kbp fragment, presumably the transposable element Tc1 (Emmons *et al.*, 1983; Liao *et al.*, 1983). The larger pDP#MM008 cross-hybridizing fragment from *unc-119(e2498)* was subcloned and shown to contain Tc1 by its ability to hybridize to a Tc1-specific probe, by a partial sequence of the insertion site, and liberation of the 1.6 kbp insertion by Eco RV, which digests in the Tc1 terminal repeats (Rosenzweig *et al.*, 1983). This Tc1 element contained a *Hind* III restriction site, which is unusual in the Bristol strain, but has been seen in Tc1 elements isolated from the Bergerac strain of *C. elegans* (Rose *et al.*, 1985), consistent with the isolation of *e2498* from a Bristol/Bergerac hybrid. Furthermore, the pDP#MM008 clone faintly cross-hybridizes to a much larger ~9 kbp *Sst* I fragment from CB1517, suggesting that pDP#MM008 contains DNA that is largely absent in this strain, and that pDP#MM008 identifies the site where breakage occurred to generate *eDf2*. At this point, we became aware of the placement of a new cosmid, M142, on the physical map in the region between T27G1 and T19F2; this cosmid contains the entire *vab-7* gene (J. Ahringer, personal communication). M142 was found to contain sequences that hybridize to pDP#MM008, and was used to extend overlapping clones to the right. The breakpoints of *eDf2* and *eDp6* were placed on the physical map by cross-hybridization to the subclones in the region, and it was determined that strain CB1517 is missing approximately 8 kbp of DNA that is present in N2 (see Figure 2.7).

In order to verify that this region also contained the *unc-119* gene, cosmid M142 DNA was injected into the gonads of *unc-119(e2498)* hermaphrodites.

This technique allows a small proportion of the progeny to inherit the transgenic DNA (Fire, 1986; Mello *et al.*, 1991). The cosmid was injected along with plasmid pRF4, which contains the dominant *su1006* allele of the collagen gene *rol-6*, allowing transformants to be scored by their characteristic rolling phenotype (Mello *et al.* , 1991). Rescue of the mutant phenotype was obtained with M142 and smaller subclones, that defined a minimum rescuing region of 5.5 kbp (see Figure 2.7). While rescue was initially only scored as restoration of wild type movement and body shape, transmitting lines were also found to be rescued for all other traits tested (see Figure 2.2). One of the smaller subclones, pDP#MM019, was able to restore partial movement to homozygous *e2498* animals, such that the *rol-6* rolling phenotype was much clearer than in *unc-119* mutants transgenic for pRF4 alone, but the animals were still partially defective in movement, and slightly dumpy, although they could form dauer larvae. This suggests that partial *unc-119* function is present in this subclone. A smaller subclone and a clone from outside the *unc-119* region (not shown in the figure) failed to rescue. Clones that were able to rescue *e2498* were also able to rescue the *ed3*, *ed4* and *ed9* alleles. When the *unc-119*-rescuing subclones were injected into animals of strain CB1517 [*eDf2; eDp6*], they were able to restore normal movement, consistent with the model that *unc-119* is the only gene affected by this chromosomal aberration.

2.3.4 unc-119 cDNAs

The partial-rescuing clone pDP#MM019 was used to screen a mixed-stage cDNA library. Three positive clones were obtained, all of which had identical 3' ends. The longest cDNA appears to be nearly complete, since it extends in the 5' direction beyond the end of the minimal rescuing region, and beyond the largest open reading frame (ORF). The size of the predicted message is 1000 b. The intron/exon arrangement of the largest cDNA is shown in Figure 2.7, and the sequence of the cDNA and genomic DNA is given in Figure 2.8. The first ATG in the cDNA does not appear to be used; it precedes both a 33 bp region that is 85% A-T rich, and the sequence 5'-TAGTTAGTTAA-3', which encodes stops in all three frames. The sequence of genomic DNA spanning the rescuing region, and extending upstream of the 5' end of the longest cDNA, was determined in parallel. Some of the genomic sequences were also determined, and communicated, by the *C. elegans* Genome Sequencing consortium (Wilson *et al.*, 1994). There is no further potential coding information in the 650 bp 5' to the first base of the longest cDNA. The predicted transcript contains five exons spanning 3.6 kbp of genomic DNA. The sequences of 5' RACE products (see Materials and Methods) contain one more base than the longest cDNA, suggesting that the longest cDNA is full-length, and that the transcript is not trans-spliced to SL1 or SL2 (Krause and Hirsh, 1987; Huang and Hirsh, 1989).

The sequences of the cDNA clones predict a single open reading frame (ORF) in the rescuing region. This ORF predicts that *unc-119* encodes a 219 amino acid protein. Search of the protein databases using the BLAST protocol (Altschul *et al.*, 1990) detected no significant similarity to other proteins in the databases, other than to C27H5.1, an unidentified ORF in *C. elegans* predicted by the genome sequencing project (GenBank Accession Number U14635; Wilson *et al.*, 1994). A 92 amino acid region shows 26% identity and 51% similarity between C27H5.1 and UNC-119. The other nematode homologue maps to chromosome *II*, to the right of the cloned genes *tra-2* and *unc-104*. Although there are several mutations that map to the general location of this

ORF, no mutant phenotype has been specifically correlated to the predicted gene.

It was not unexpected to find a cDNA that extended beyond pDP#MM019, which only contains Exons III to V, since phenotypic rescue was only partial with this clone. This rescue, which occurs without the presumed *unc-119* promoter, may be possible through low-level constitutive expression of the genes in the extrachromosomal DNA. Transgenic DNA frequently undergoes extensive rearrangement and is present in many copies as a single array (Mello *et al.*, 1991). There is an in-frame ATG at the start of Exon III, that may permit translational initiation of a truncated form of UNC-119 to produce a polypeptide containing the carboxy terminal 196 of 219 amino acids.

In order to verify that *unc-119* had been cloned, oligonucleotide primers were designed that would allow PCR amplification of the majority of the coding region (see Materials and Methods). Genomic DNA from *ed3*, *ed4, ed9* and the revertant allele *e2498ed5* was PCR amplified and the products were subcloned. This allowed sequencing of 25 bp of Exon III and all of Exon IV. Both of the EMS-induced alleles *ed3* and *ed4* contain single C to T transitions in Exon IV that create in-frame stop codons. The allele *ed9* contains a G to A transition in the intron just preceding Exon IV, that changes the splice acceptor sequence from 5'-TTTCAG-3' to 5'-TTTCAA-3'. Since intron splice acceptors in *C. elegans* are invariant for the last two bases of this sequence (Wood, 1988), this change is expected to disrupt normal processing of the primary transcript. The revertant *e2498ed5* appears to be the result of a perfect excision event of Tc1, as no base changes from wild type were found. Since we did not check the sequences of *ed3*, *ed4* and *ed9* upstream of the last 25 bp of Exon III, we cannot rule out additional base changes in the upstream regions in these mutants.

To test whether the cDNA sequences were sufficient to rescue the phenotype, the longest cDNA clone was fused at the *Ase* I site in the first exon with a 1.2 kbp genomic DNA fragment containing the upstream region to the *Hind* III site (Figure 2.7). Arrays carrying this 'mini-gene' (pDP#MM051) were able to restore all aspects of the mutant phenotype in the F_1 progeny when injected into strains homozygous for each of the mutant alleles of *unc-119* (rescue of *ed3* is shown in Figure 2.2), consistent with the requirement of only one mRNA from this region for *unc-119* function.

2.3.5 Expression of unc-119

In order to determine the pattern of expression of *unc-119*, a fragment containing the 5' half of the *unc-119* coding region and 1 kbp of DNA upstream of the 5' end of the longest cDNA was fused in-frame with the *lacZ* gene encoding *E. coli* β-galactosidase, using the *C. elegans lacZ* expression vector pPD22.04 (Fire *et al.*, 1990). The portion of *unc-119* in the fusion construct does not encode enough of the protein to rescue, since the mutations in *ed3* and *ed4* are further toward the carboxy-terminus than the fusion junction. The *unc-119::lacZ* fusion construct was transformed into wild-type animals along with the dominant *rol-6* marker. Animals from mixed-stage plates of a strain with approximately 50% meiotic transmission (as determined by the presence of the 'Rol' phenotype in the progeny) were fixed and stained with both DAPI and X-gal.

The fusion enables staining of many neurons in larval and adult animals (see Figure 2.9). The transgene is present as an extrachromosomal array, and is sometimes lost during mitosis, resulting in a number of animals that are mosaic for the transgene, and stain in only a subset of cells. Although the *unc-119::lacZ* fusion contains a nuclear-localizing signal (NLS; see Fire *et al.*, 1990), several hours after fixation and treatment with X-gal, staining begins to appear in axons and process bundles. Staining is seen in embryos with as few as 60 cells based on the counting of DAPI-stained nuclei. By the time development has reached 'comma' stage, about 400 minutes after fertilization, staining appears localized largely to the anterior portion of the embryo (Figure 2.9b). After hatching, and continuing through larval and adult development, staining is visible in both the dorsal and ventral nerve cords and the nerve ring (Figure 2.9, a and e). Cell bodies of neurons which innervate the vulva can be seen in Figure 2.9c. Differences in staining can be seen near the tail between males and hermaphrodites, consistent with expression in lumbar ganglion cells (Figure 2.9d), where particular neurons develop to innervate the tail rays in the male (Sulston *et al.*, 1980). In the hermaphrodite tail, staining is seen that strongly resembles the positions of neuronal cell bodies in the pre-anal and lumbar ganglia (Figure 2.9f; Siddiqui and Culotti, 1991). There is also some evidence of expression outside of the nervous system; in Figure 2.9a, staining can be seen anterior of the anterior pharyngeal bulb, where no neuronal cell bodies are found (White *et al.*, 1986).

2.4 Discussion

We have identified a new gene, *unc-119*, that appears to be involved in neuron function or development in *C. elegans*. The DNA encoding UNC-119 has been identified based on several sets of evidence. First, homozygous *unc-119* animals carrying the presumptive coding region (either an intact genomic DNA fragment, or a promoter:cDNA chimera) as a transgene are phenotypically normal. Second, strains that demonstrate the *unc-119* mutant phenotype contain sequence changes that are predicted to cause loss-of-function mutations in the predicted protein (*ed3* and *ed4*) or impede the production of a normal message (*e2498* and *ed9*). We have also shown that *unc-119* is nearly completely deleted in *eDf2; eDp6*, and is likely the only gene affected. Therefore, an intact coding region is necessary, and sufficient, for UNC-119 function.

The observation of partial phenotypic rescue by a truncated *unc-119* gene suggests that specific promoter or expression sequences that may lie upstream of the gene are not essential. It is possible that rescue is accomplished by low level constitutive expression from many tandem copies. Partial rescue of the *unc-51* mutant phenotype by transgenic DNA containing only minimal upstream sequences has been reported (Ogura *et al.*, 1993), but in that case the coding region was supposedly intact. The protein sequence predicted from the pDP#MM019 subclone is missing the first in-frame ATG, suggesting that either this ATG is not used for initiation *in vivo*, or the first 20 amino acids of UNC-119 are not essential for function, and that partial rescue with pDP#MM019 is due to lower amounts of expression.

The *unc-119* mutant phenotype is most apparent in its lack of movement, as mutants lack coordinated locomotion from hatching through adulthood. There is indirect evidence that this defect results from a disruption in neuron placement or function. Examination of muscle by EM, polarized-light microscopy and antibodies to thin filaments show no structural defects, and since mutants can hypercontract after exposure to levamisole, at least partial muscle function remains.

More compelling evidence of a role for UNC-119 in the nervous system comes from the defective ability of homozygous *unc-119* animals to respond to chemical signals in their environment. Although *unc-119* animals can move, albeit slowly, they do not appear to recognize a change in the abundance of food, as their egg laying, locomotory and pharyngeal responses do not change appropriately. This is best reflected in the inability of *unc-119* animals to form dauer larvae, similar to what is seen in *daf-10* mutants, which are dauer-defective (Daf-d) due to a chemosensory abnormality (Albert *et al.*, 1981). This is the opposite phenotype of dauer-constitutive (Daf-c) mutants such as *daf-7* and *daf-11*, which inappropriately enter dauer in the presence of food (Riddle *et al.*, 1981). The observation that *unc-119* can suppress the Daf-c phenotype of *daf-11* but not *daf-7* is consistent with the placement of *unc-119* in the pathway for dauer formation, at a position downstream of *daf-11* but independent of *daf-7*, at the same step as *daf-10* and other genes implicated in chemosensation (Thomas *et al* ., 1993). This suggests that mutants have an inability to recognize the absence of food, which would also explain the maintenance of pharyngeal pumping on removal from the bacterial lawn. Since a defect in chemosensation alone is insufficient to explain the mobility defect in *unc-119* animals (Albert *et al.*, 1981), lack of UNC-119 may affect other aspects of the nervous system. *unc-31*, for example, while having a locomotory defect, also shows constitutive pharyngeal pumping and defects in egg-laying and dauer larvae response, although those defects are different than those seen for *unc-119* (Avery *et al.*, 1993).

Evidence for a direct role in neuron development or function, however, comes from the *in vivo* expression pattern of a reporter gene construct. The 5' end of *unc-119* fused to the *E. coli lacZ* gene allows the X-gal staining of many neurons. This includes those whose function is not locomotory, such as those innervating the male tail and hermaphrodite vulva. The apparent neuronal defects seen in *unc-119* mutants are consistent with the expression patterns as seen with the *lacZ* fusions. This consistency is both temporal (gene expression beginning in the embryo, and movement defects are seen in L1 larvae immediately after hatching) and spatial (effects on dauer formation and movement consistent with expression in the sensory and motor nervous systems).

If UNC-119 is expressed in a large number of neurons, why is the phenotype of *unc-119* null mutants relatively mild? Perhaps UNC-119 is part of a larger gene family, with redundant function. The similarity of *unc-119* to a predicted ORF detected elsewhere in the genome by the genome sequencing project of *C. elegans* supports the idea that other genes may contribute to the same function. However, there is also precedent for a weak phenotype for mutations in well-conserved neuronal proteins. Synaptotagmin is a conserved transmembrane protein expressed in the *C. elegans* nervous system as an abundant protein which appears to be localized in synaptic vesicles (Nonet *et al.*, 1993). *C. elegans* synaptotagmin is over 70% similar to vertebrate synaptotagmin, and unlike in vertebrates, there appears to be only one gene in nematodes (*snt-1*). However, *snt-1* null mutants are viable, although slow growing, and uncoordinated (Nonet *et al.*, 1993). More recently, Ségalat *et al* . (1995) and Mendel *et al* . (1995) report the phenotype of worms carrying loss-of-function mutations in the Gα_0 subunit of the heterotrimeric G-protein. Although this protein is apparently expressed in every neuron in *C. elegans*, in addition to other, non-neuronal cells, and has been thought to play an important role in neuronal development, loss-of function mutants display a surprisingly subtle phenotype.

A specific role in neuronal development or function is not yet obvious, as the subcellular localization of UNC-119 has not been determined, and a non-nematode counterpart has yet to be identified. Since some Unc genes first cloned in *C. elegans* turn out to have important homologues in other systems, it is possible that *unc-119* may have a fundamentally conserved function in higher organisms.

The expression of UNC-119 may also have use as a general marker for the *C. elegans* nervous system. There have been several reagents reported which allow observation of all (Hekimi, 1990) or a subset of the nervous system in *C. elegans* (Hedgecock *et al.*, 1985; Hope, 1991; Siddiqui and Culotti, 1991; Hamelin *et al.*, 1992; McIntire *et al.*, 1993; Miller *et al.*, 1993; Nonet *et al.*, 1993; Chen and Lim, 1994), yet only *unc-51* has been reported to allow observation of the nervous system in the developing embryo (Ogura *et al.*, 1994). It is possible that *unc-119*::reporter fusions may allow the specific visualization of the process of nervous system development in *C. elegans*, and hence facilitate the recovery and characterization of mutations affecting morphology, number, placement or migration of neurons.

Acknowledgements

We thank R. Bhatnagar for technical assistance with electron microscopy and ultramicrotomy; M. Gilbert and D. Moerman for help with antibody staining and for donating the anti-vinculin antibody; A. Fire for donating the *lacZ* vectors; B. Barstead for the λZAP cDNA library; J. Hodgkin for strains and advice and T. Barnes for the dosage-blotting idea. Some strains used in this work were obtained from the Caenorhabditis Genetics Center, which is supported by a contract between the NIH National Center for Research Resources and the University of Minnesota. Some of the genomic sequences in the *unc-119* region were kindly communicated prior to general release by the *C. elegans* Genome

Sequencing consortium. This work was supported by grants from the Natural Sciences and Engineering Research Council and the Alberta Heritage Fund for Medical Research. The initial identification and characterization of *unc-119* by D.P. took place at the MRC Laboratory of Molecular Biology, Cambridge, UK, where he was supported by a Postdoctoral Research Fellowship from the Human Frontiers Science Program.

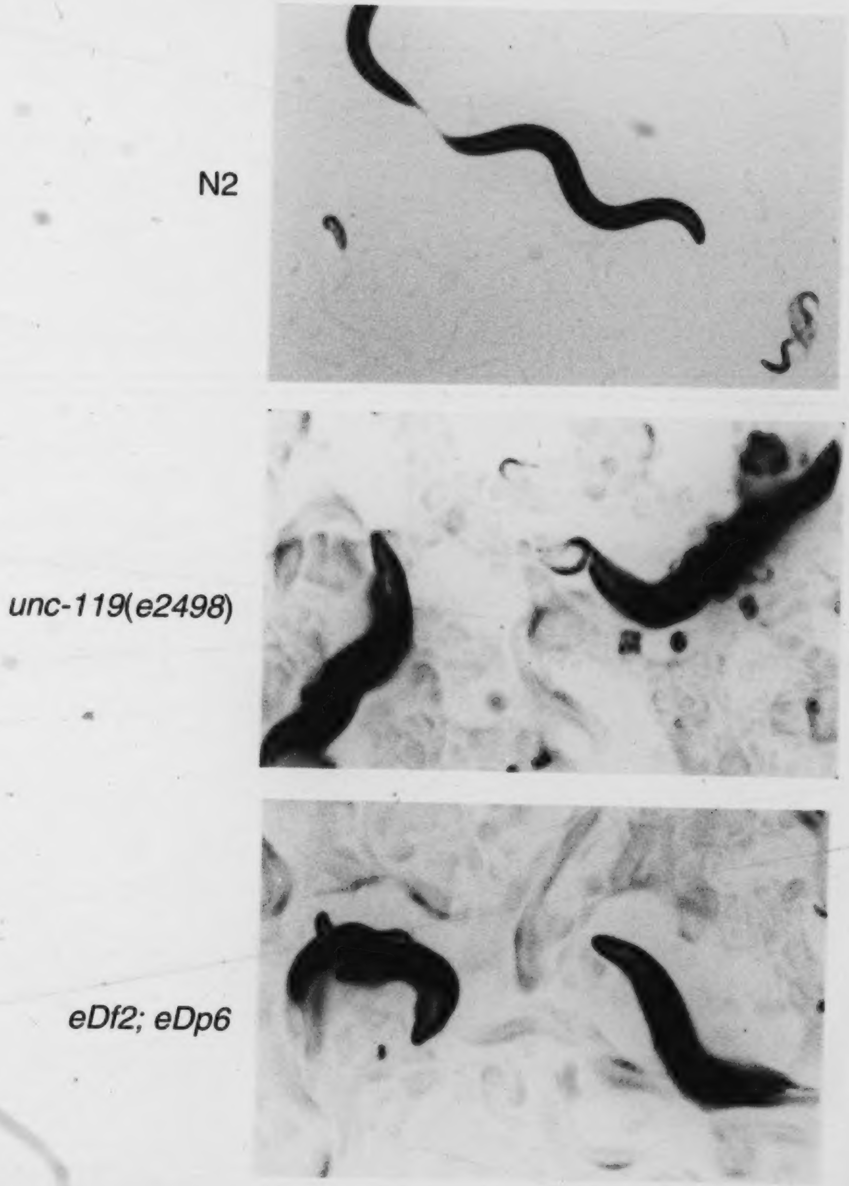


Figure 2.1 Photographs of wild-type (N2), *unc-119(e2498)* and *eDf2; eDp6* animals. The top panel shows an L4 animal displaying the characteristic sinusoidal wave during motion. The other panels show the similarity in *unc-119(e2498)* and *eDf2; eDp6* adults. The mutant animals are severely uncoordinated and tend to curl. Mutants also show a slight 'dumpy' morphology. All photographs are at the same magnification; the L4 animal in the top panel is approximately one mm long.

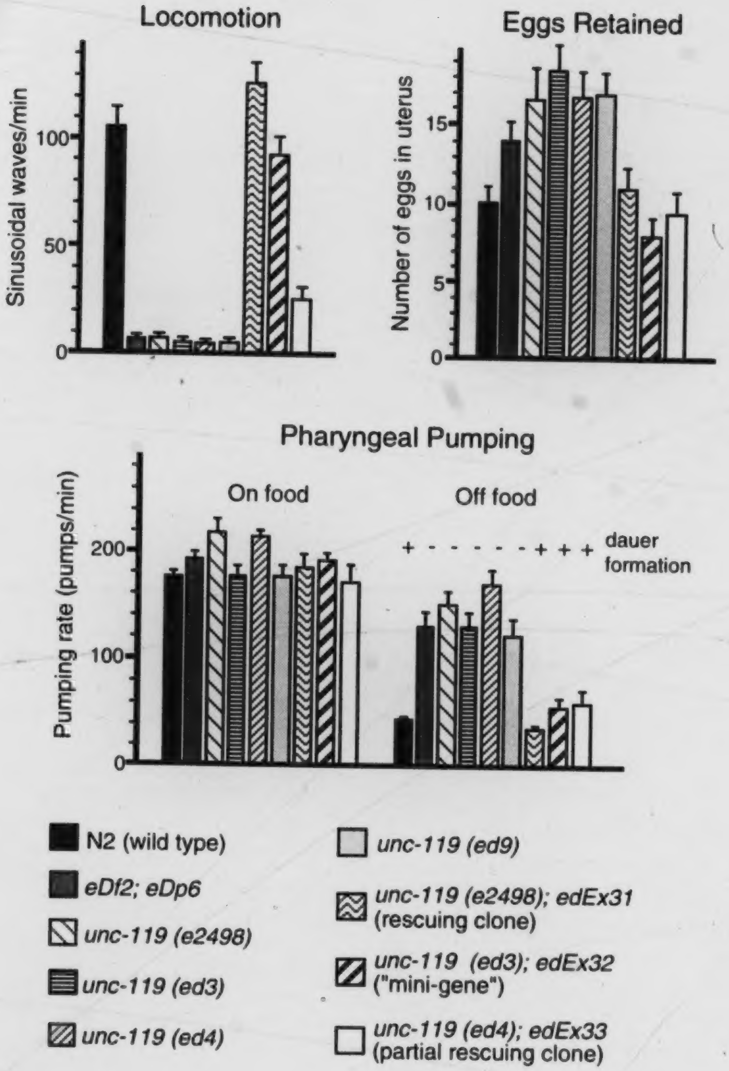


Figure 2.2 Quantitation of defects in wild type, *unc-119* homozygous mutants and homozygous *unc-119* animals carrying the putative *unc-119* wild-type gene (or portions thereof) on transgenic arrays. In the third histogram, dauer-forming ability is indicated above the 'off food' bars as either '+' or '-'. Bars show the standard error of the mean (SEM). Data collection is described in Materials and Methods.

2

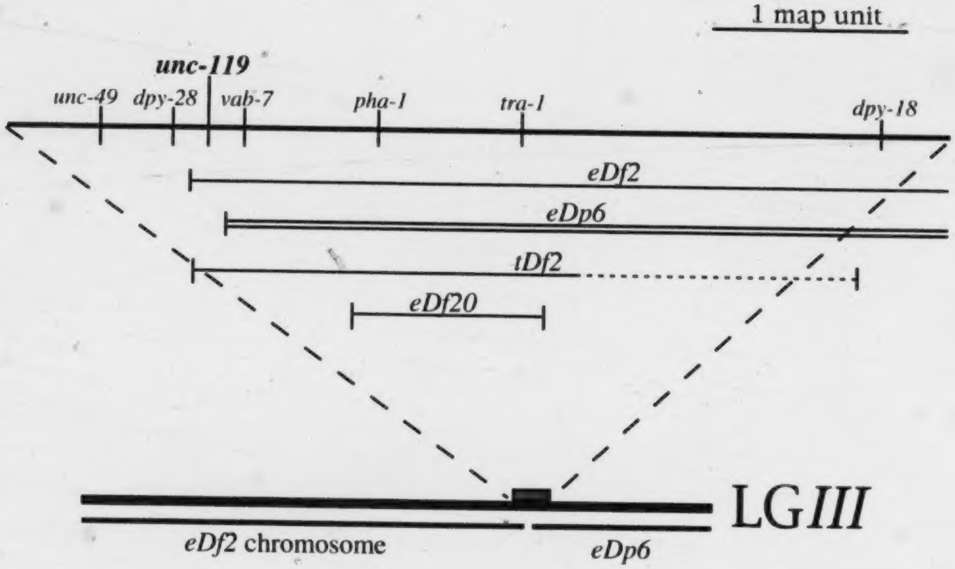


Figure 2.3 Genetic map of a portion of chromosome *III* (right end), expanded to show the *unc-49* to *dpy-18* interval (Edgley and Riddle, 1993). A subset of the genetic markers in the region are shown. Chromosomal rearrangements are indicated below the genetic map, deficiencies as single lines, and duplications as double lines. The left and right breakpoints of *eDf20* have been physically mapped relative to *pha-1* and *tra-1* (Hodgkin, 1993; Granato *et al.*, 1994). As is discussed in the text, the portion of chromosome *III* retained in *eDf2* is shown as '*eDf2* chromosome.'

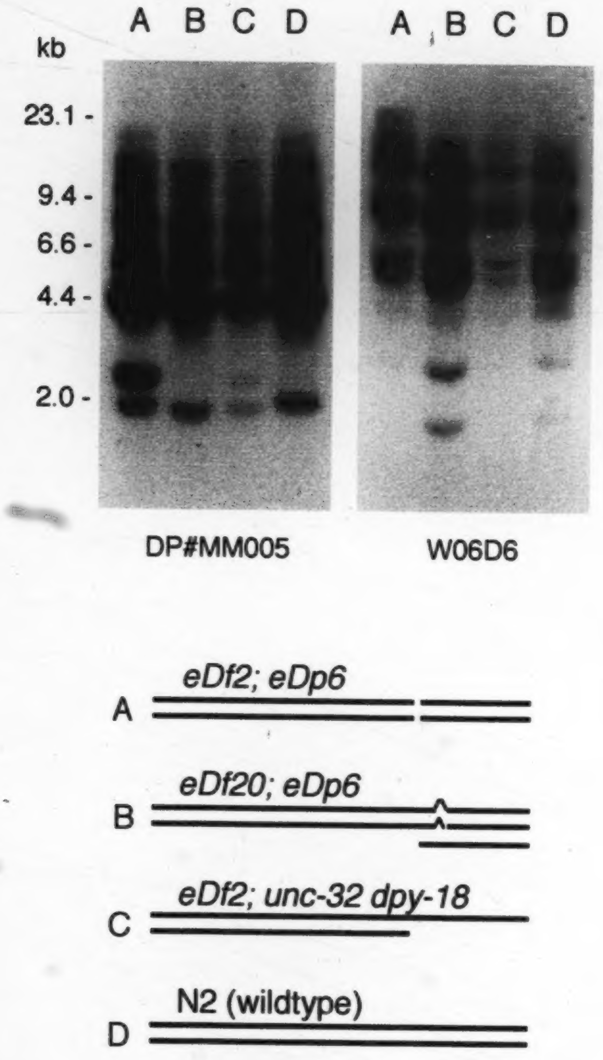


Figure 2.4 Southern analysis of *Hind* III digests of strains containing varying copy numbers of rearrangements affecting the right end of chromosome *III*, and representation of chromosome *III* for each of the strains used. The same Southern blot was probed with two different clones, pDP#MM005, which is contained on the *eDf2* chromosome, and W06D6, which is contained on *eDp6*. In the right hand photograph, much more intense cross-hybridization is seen in lane B than in lane C due to the presence of more *eDp6* DNA. Approximate sizes (in kbp) are shown on the left.

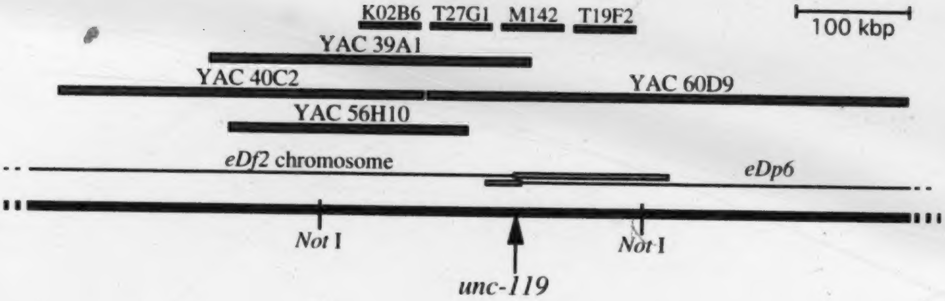


Figure 2.5 Physical map of the *unc-119* region, showing Yeast Artificial Chromosome (YAC), and cosmid clones. The physical map has the same alignment as the genetic map in Figure 2.3. *Not*I restriction enzyme recognition sites are indicated. The alignment was determined by YAC:cosmid cross hybridization, as well as long-range restriction mapping. The open boxes on the end of *eDf2* and *eDp6* indicate the presence of DNA, suggested by the size of the *Not*I fragment on pulsed-field gels, which fails to cross-hybridize with the cosmid clones directly above. YAC clone 39A1 has been previously shown to hybridize to both *eDf2* and *eDp6* DNA (Albertson, 1993). The boxes below the cosmid names do not indicate their size, but are meant solely to show the region of cross-hybridization. The cosmids W06D6, W03F10, M02H11 and W09D6, which map to the right of *unc-119* and only cross-hybridize to Y60D9 (see text) are not shown on this figure since they were not tested for cross-hybridization to the 300 kbp *Not*I fragment.

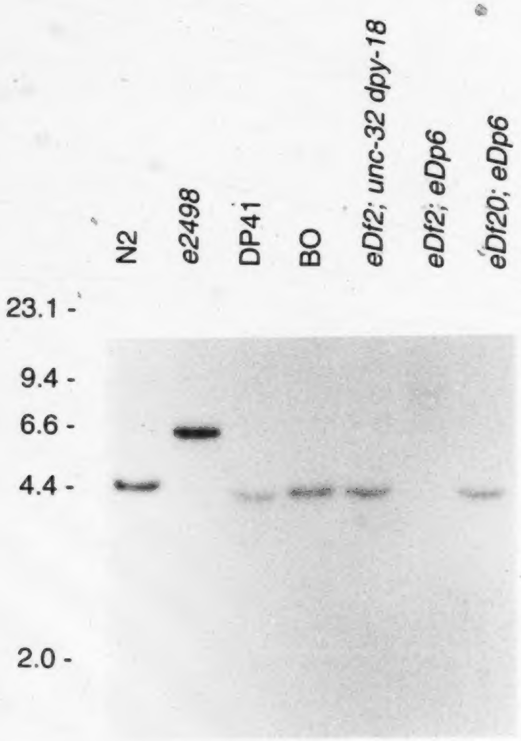


Figure 2.6 Southern blots of *SstI* genomic digests of N2, BO, *unc-119(e2498)*, CB1517 [*eDf2; eDp6*] and the revertant DP41, probed with pDP#MM008. The common band in N2, DP41, BO, *eDf2; unc-32 dpy-18* and *eDf20; eDp6* is approximately 4.5 kbp in size. The larger band in *e2498* is due to the insertion of the 1.6 kbp Tc1 transposable element. There is a larger, faint band in the *eDf2; eDp6* lane, consistent with a rearrangement of DNA in pDP#MM008, and the absence of DNA from this clone.

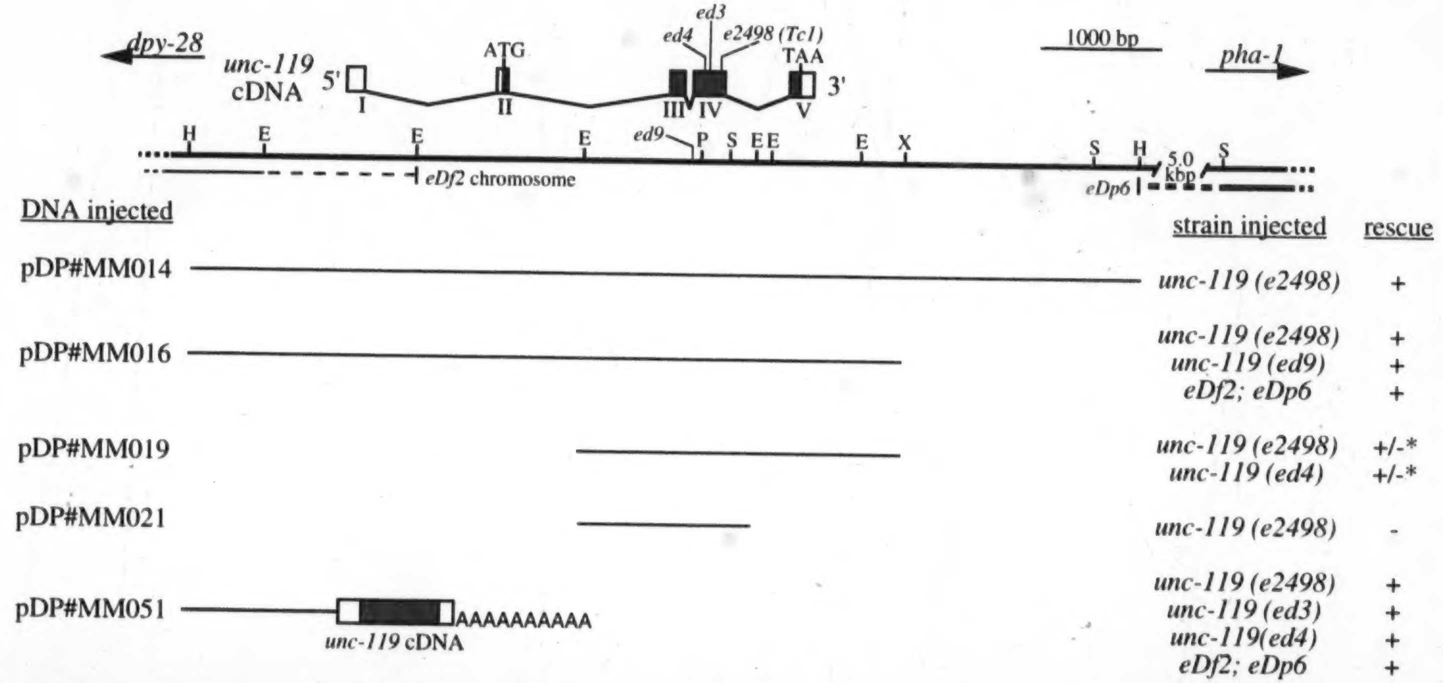


Figure 2.7 Short-range physical map showing the genomic DNA in the *unc-119* region. The cDNA is shown above the genomic restriction map (the ORF is shown as shaded), while the limits of the breakpoints of the *eDf2* chromosome and *eDp6* are shown below. The rescuing ability of various clones from the region is shown. '+' denotes complete rescue of the mutant phenotype, '+/-' denotes partial rescue, and '-' indicates failure to rescue (see Figure 2.2). The pDP#MM051 clone is a 'minigene' constructed by ligating the upstream genomic DNA to the full-length cDNA at an *Ase*I site in Exon I (see text). The sites of the Tc1 insertion in *e2498* and the nonsense mutations in *ed3* and *ed4* are shown in the cDNA. The site of the change in *ed9* is shown above the genomic DNA. The fusion to the *lacZ* vector pPD22.04 was made at the *Pst*I site (see text). H=*Hin*dIII, E=*Eco*RI, S=*Sst*I, P=*Pst*I, X=*Xba*I. *Eco*RI and *Pst*I sites to the right of the *Xba*I site have not been determined and so are not shown.

```
gaattctatgaaaatgtctaaagaaaatggggaaacaatttcaaaaaggcacagtttcaatggtttccgaattatactaaatccctctaaaaacttccgg 100
caaattgatatccgtaaaagagcaaatccgcattttgccgaaaattaaaatttccgacaaatcggcaaaccggcaatttggcgaaatttgccggaacga 200
ttgccgcccacccctgttccagaggttcaaactggtagcaaagctcaaaatttctcaaattctccaatttttttttgaattttggcagtgtaccaaaatg 300
acattcagtcatattggtttattatagatttatttagataaaatcctaaatgattctacctttaaagatgcccactttaaaagtaatgactcaaacttca 400
aattgctctaagattctattgaattaccatctttcctctcatttctctcactgtctatttcatcacaaattcatccctctctcctctcttctctctcc 500
ctctctctctttctctttgctcatcatctgtcatttgtccgttcctctctctgcgccctcagcgttccccacactctctcgcttctcttttcctaga 600
cgtcttctttttttcatcttcttcagcctttttcgccatTTTCCATCTCTGTCAATCATTACGGACGACCCCCATTAATTTTTGGATGTGCTCTGCGAGCA 700
AGAGCACACGGAACTCATTTATTTTTTGTTGGAATTTCTTTTTCATTTTAGTTAGTTAAAAGgttattcagtgcaaaaattgatttttttgctcttttggt 800
acattagaggggtttgaaagttttccataaaaagggtttccggagtaaaaaatctaaaaattttacaaacatttccaaaaattttgggaacatatcttt 900
tcaaaaactcaaaaaaaaacgcaaacatcataaatttaatcaattaattggatgcgttcaaatattttaaaaatatttttgaattgcatatttctttact 1000
cataaaaaaattattaatgtttaagtaaaaaaaatatttcaggatttaaacaaaaactgtaaatcccataaaacagtgttattttttagttatttttag 1100
taattatttgtgaattaaaaacacatttttttcaaaatattccttgaataaataaaattgaaagtgaattcaaaatatttcaccaattttacaaaattt 1200
gaaaatttgaaaaaatatttctaccaaattgtgctcaaaaaatattattaagttcccaggaaatttatgatgaaaacttggaatatttttggtaaaatt 1300
ttctcaaaaaattgaaagttgtcaaaaacacactaacactatacctcgcggcatagaaaaaactggtggccgaatttttaaaccaatttaaattaaata 1400
tatttttgcactttgtacaaaaaaatttgaagttcccaggaaattcatacggaaatatttccagaactttataatttttattttaattctcaaaattt 1500
acaataaaaacttggaaacctgaaatttgcttaaagaataaccctatcccccacggcctagaaaatactggtggccgaatttttccgcagccacaccac 1600
                                                                      M  K  A  E  Q  Q  Q  Q  S  I  A  P
ctctaatctccagagcactctccaaaattcccataatccccaaaattttcagCTACAACAGCATATATGAAGGCAGAGCAACAACAACAATCGATCGCAC 1700
  G  S  A  T  F  P  S  Q
CCGGCTCGGCAACCTTCCCGTCTCAGgtgaggaagacgcagagaaatgagagagagagcgtgtttatgcacgcagtatactaactcctgatatccaattt 1800
catatatcgatatgcttgctttcttttcttctacgaagaaattgtttgaacttttctcaaaaaacgcggatttagagagacttagagttgctccaaata 1900
aatttaatttgcaattttgcacttactattagaatctactttatagcgctaaaatgaccgaataacacagtagatttaaaagaaaatttaaaaatta 2000
tttttttgtgcagtcaaatttcgctagttttgccactcatttttgaacatttgtatatgtgcctttcggtggttcaatcaggggtggcacgtggtgtca 2100
gagtgttccatttcggtttgatctacgttgatctacaaaaaatgcgggagaagagacgcagagttctcaactgatttcacatggttatgaaccaggggtg 2200
gttggcaaatttgccgtttgccgagctgggcaaattttgagatttgccgcacaccccttttatgaacgtgctgacgtcacatctttttggacaaagaatt 2300
cccgcatttttgtatatcaaaccgtgatgggacagcctggcaccacgtggggttggcggcaaatgatttttccgacaaatcggctttaaggcaaattg 2400
ccggaacccaaagtttctggcaaattgttgatttgccgaaaaaaatttcaaaaaggcacagttttaagtgtttccgtcttataaaatatttctctaaaa 2500
```

Figure 2.8 Nucleotide sequence and predicted amino acid sequence of the *unc-119* locus. cDNA sequences are shown in capital letters, and the predicted amino acid sequence is denoted by single-letter abbreviations above the cDNA sequence. The first (unused) and second ATG are underlined, as are the stop codon and the putative polyadenylation site. The sequence has been submitted into GenBank (Accession Number U32854). The sites of the mutations are indicated as follows: *=*ed9* [splice acceptor sequence 5'-...AG-3' → 5'-...AA-3']; † = *ed4* [CAA(gln) →TAA(ochre)]; ‡= *ed3* [CGA(arg)→TGA(opal)]; ↓= *e2498* [Tc1(Hin) insertion 5'-...GAACA(Tc1)TATCT...-3']. (continued on next page)

```
atttccggaaaacggaaaattgccggaaatacaaagtttcgtcaaatcggcaaacctgcaatttgccgaaccaactgccgccaacacctggttcccaaa 2600
aaagctgcactgtttgaatattaaaaaaaaaattctctcaaaaataataaattgtgaaataggctaaattcacatctatgtacaaaaattgcatgccagc 2700

                                                                                               M  P
accggtcaattttctcaattactttttgtctattttgccgaacataatttgctgtttccttggcagaatacagaaaactcttttttttttcagATGCCC 2800
R  P  P  P  V  T  E  Q  A  I  T  T  E  A  E  L  L  A  K  N  Q  I  T  P  N  D  V  L  A  L  P  G  I  T
CGACCCCCACCGGTAACCGAACAGGCTATAACCACCGAGGCGGAGCTTCTCGCGAAAAATCAAATTACACCAAATGATGTGCTAGCTCTTCCAGGAATCA 2900
   Q\  G
                                                                           *  F  L  C  S  P  S  A  N
CTCAAGgtaagtgaatatttgctctattgataaacgcgccgtgtactccacgtggacaaacatatcaaattttcagGATTCCTATGCTCCCCATCGGCAA 3000
  V  Y  N  I  E  F  T  K  F  Q  I  R  D  L  D  T  E  H  V  L  F  E  I  A  K  P  E  N  E  T  E  E  N
ACGTCTATAACATTGAGTTCACCAAATTCCAAATCCGTGACCTCGACACGGAACACGTGCTCTTCGAAATTGCAAAACCGGAAAATGAGACGGAAGAGAA 3100
 L  Q  A  Q† A  E  S  A  R  Y  V  R  Y  R‡ F  A  P  N  F  L  K  L  K  T  V  G  A  T  V  E  F  K  V
TCTGCAGGCGCAAGCCGAATCGGCAAGATATGTCCGATATCGATTTGCGCCGAATTTTCTGAAATTAAAGACGGTCGGCGCGACGGTCGAGTTCAAGGTC 3200
G  D  V  P  I  T  H  F  R  M  I  E  R  H  F  F  K  D  R  L  L  K  C  F  D  F  E  F  G  F  C  M  P  N
GGCGATGTGCCGATCACACATTTTCGAATGATTGAACGGCACTTTTTCAAGGATCGTTTATTGAAATGCTTTGATTTTGAATTTGGATTCTGTATGCCAA 3300
  S  R  N  N  C  E  H ↓I  Y  E  F  P  Q  L  S  Q  Q  L  M
ATTCGAGAAACAACTGTGAACATATCTATGAATTTCCACAACTTTCACAACAACTCAgtgagtttgaagcgggaaaaattatttgagctcattttgtcaa 3400
tagagcgaatttctaggtttttaaaaataaaataaatattgacaaaaagctgccctacagtagttcaaaatgtgaaaacttggaggaaattcgagtttt 3500
ctctgagaagtttttggcgggtaattcaaattttctgagaaaaaaattagccgcgagaattcaaattttctgagaaaattatttgacgggaaattcaaaa 3600
tttctcggaaatgtattgaacatttgagaaattttgttaaaaactacattaaaccgataaaaaagttcgcaaattttttttaagtgctcatggtgcatcg 3700
aagagaagaattcttgccaaaactagcccatgaaattaaaaattatcgcggaaaaaaactgactaattggaaagtaaaatgaaatttgcaaaaaaaaaat 3800
                                              D  D  M  I  N  N  P  N  E  T  R  S  D  S  F  Y  F  V
atgtttttttgttgtaaaattaaaaaaaaaactgtatatttcagTGGACGACATGATCAATAATCCAAACGAGACCCGCTCCGATAGCTTCTATTTTGT 3900
 E  N  K  L  V  M  H  N  K  A  D  Y  S  Y  D  A
GGAGAATAAGCTCGTAATGCACAATAAGGCCGACTACTCGTATGATGCATAATTTCCCGCCGATGGAATAAAACTAATAATTTTTCTCGCAATTTATAGA 4000
AATCATGTGATTTTCCTACCGCTAAATTTAAAATTTCCTAATTTTTCACTATATTTCTCACAAAATGCATTCCGAAAACTCTCATTTTTCAGCTTTTCAA 4100
TACAAATCCCAATAAATGATTTTTCAAAtataatgttctcaattgatatgttttatttgaatgttggaaaaaaccataaaattaaaaaaaacttgaatta 4200
ttcaaaaataattcggccctttacaaaaaattatcataaaaattaaatttggaaaatctgtgatttgtcgctttgccggaagtttccaattccggcaac 4300
taaccgatttgccgattagctagaaatttcaattccgacaatgtgccgctttgccgatttgccggaaatttccaatttaccggcaattaacggacttgcc 4400
gattagctagaaatttaactttagataaacttctcaacaaattcaaattggaaagttcagcgatttgccgcttgaccggaagtttcagttaccggacttc 4500
agccattgcgattgcaaagtttactgcaaatgcgcatg-3' 4538
```

Figure 2.8, continued

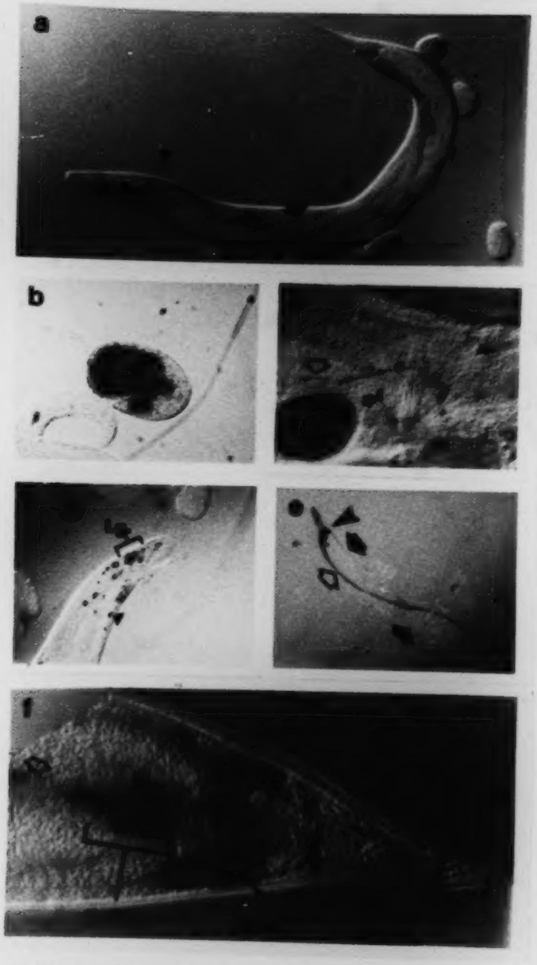


Figure 2.9 X-gal staining of fixed *unc-119::lacZ* transgenic animals. (a) Adult hermaphrodite, showing ventral nerve cord cell bodies, the dorsal nerve cord and the nerve ring. Anterior to the nerve ring, staining can be seen in the pharynx. (b) Comma stage embryo. (c) Ventral nerve cord near the hermaphrodite vulva. An asterisk (*) indicates a VC neuron cell body, above and to the right of a developing embryo. (d) Adult male tail showing cell bodies in the posterior part of the lumbar ganglion. (e) Adult hermaphrodite, showing staining in the dorsal and ventral nerve cords and the nerve ring. (f) Magnification of posterior part of animal in (e), a ventral view showing the pre-anal and lumbar ganglia. All animals result from a strain carrying the transgene extrachromosomally, and may be mosaic for the transgene due to mitotic loss. ⇨, ventral nerve cord; ➡, dorsal nerve cord; ▲, cell bodies in ventral nerve cord; ▲, nerve ring; l.g., lumbar ganglion; p.a.g., pre-anal ganglion.

2.5 Bibliography

Albert, P. S., Brown, S. J., and Riddle, D. L. (1981). Sensory control of dauer larvae formation in *Caenorhabditis elegans*. J. Comp. Neur. *198*, 435-451.

Albertson, D. (1993). Mapping chromosomal rearrangement breakpoints to the physical map of *Caenorhabditis elegans* by fluorescent *in situ* hybridization. Genetics *134*, 211-219.

Albertson, D. G., and Thomson, J. N. (1976). The pharynx of *Caenorhabditis elegans*. Philos. Trans. R. Soc. Lond. B Biol. Sci. *275*, 299-325.

Albertson, D. G., and Thomson, J. N. (1982). The kinetochores of *Caenorhabditis elegans*. Chromosoma *86*, 409-428.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. J. Mol. Biol. **215**:403-410.

Anderson, P. (1989). Molecular genetics of nematode muscle. Ann. Rev. Genet. *23*, 507-255.

Avery, L., and Horvitz, H. R. (1990). Effects of starvation and neuroactive drugs on feeding in *Caenorhabditis elegans*. J. Exp. Zool. *253*, 263-270.

Avery, L. (1993). The genetics of feeding in *Caenorhabditis elegans*. Genetics *133*, 897-917.

Avery, L., Bargmann, C. I., and Horvitz, H. R. (1993). The *Caenorhabditis elegans unc-31* gene affects multiple nervous system-controlled functions. Genetics *134*, 455-464.

Barstead, B., and Waterston, B. (1989). The Basal Component of the Nematode Dense-body Is Vinculin. J. Biol. Chem. *264*, 10177-10185.

Birren, B. and Lai, E. (1993). *Pulsed Field Gel Electrophoresis: A practical guide.* Academic Press Inc., San Diego, CA.

Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. Genetics *77*, 71-94.

Cassada, R. C., and Russell, R. L. (1975). The dauer larvae, a post-embryonic developmental variant of the nematode *Caenorhabditis elegans*. Dev. Biol. *46*, 326-342.

Chalfie, M., and White, J. (1988). The Nervous System in *The Nematode Caenorhabditis elegans*, edited by W.B. Wood. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Chen, W., and Lim, L. (1994). The *Caenorhabditis elegans* small GTP-binding protein RhoA is enriched in the nerve ring and sensory neurons during larval development. J. Biol. Chem. *269*, 32394-32404.

Coulson, A., Sulston, J., Brenner, S., and Karn, J. (1986). Toward a physical map of the genome of the nematode *C. elegans* PNAS (USA) *83*, 7821-7825.

Coulson, A., Waterston, R., Kiff, J., Sulston, J., and Kohara, Y. (1988). Genome linking with yeast artificial chromosomes. Nature *335*, 184-186.

Coulson, A., Kozono, Y., Lutterbach, B., Shownkeen, R., Sulston, J., and Waterston, R. (1991). YACs and the *C. elegans* genome. Bioessays *13*, 413-417.

Edgley, M. L., and Riddle, D. L. (1993) The nematode *Caenorhabditis elegans*, in "Genetic Maps: Locus Maps of Complex Genomes, Sixth Edition, No. 3, Lower Eukaryotes." edited by S. J. O'Brien. Cold Spring Harbor Laboratory Press, Plainview, NY.

Emmons, S. W., Yesner, L., Ruan, K. -S., and Katzenberg, D. (1983). Evidence for a transposon in *Caenorhabditis elegans*. Cell *32*, 55-65.

Epstein, H. F., and Thomson, J. N. (1974). Temperature-sensitive mutation affecting myofilament assembly in *Caenorhabditis elegans*. Nature *250*, 579-580.

Fire, A. (1986). Integrative transformation of *Caenorhabditis elegans*. EMBO J. *5*, 2673-2680.

Fire, A., Harrison, S. W., and Dixon, D. (1990). A modular set of *lacZ* fusion vectors for studying gene expression in *Caenorhabditis elegans*. Gene *93*, 189-198.

Frohman, M. A., Dush, M. K., and Martin, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. PNAS (USA) *85*, 8998-9002.

Granato, M., Schnabel, H., and Schnabel, R. (1994). Genesis of an organ: molecular analysis of the *pha-1* gene. Development *120*, 3005-3017.

Hall, D. H., and Hedgecock, E. M. (1991). Kinesin-related gene *unc-104* is required for axonal transport of synaptic vesicles in *C. elegans*. Cell *65*, 837-847.

Hata, Y., Slaughter, C. A., and Südhof, T. C. (1993). Synaptic vesicle fusion complex contains *unc-18* homolog bound to syntaxin. Nature *366*, 347-351.

Hedgecock, E. M., Culotti, J. G., Thomson, J. N., and Perkins, L. A. (1985). Axonal guidance mutants of *Caenorhabditis elegans* identified by filling sensory neurons with fluorescein dyes. Dev. Biol. *111*, 158-170.

Hedgecock, E. M., Culotti, J. G., and Hall, D. H. (1990). The *unc-5*, *unc-6*, and *unc-40* genes guide circumferential migrations of pioneer axons and mesodermal cells on the epidermis in *C. elegans*. Neuron *2*, 61-85.

Hekimi, S. (1990). A neuron-specific antigen in *C. elegans* allows visualization of the entire nervous system. Neuron *4*, 855-865.

Hodgkin, J. A. (1980). More sex determination mutants of *Caenorhabditis elegans*. Genetics *96*, 649-664.

Hodgkin, J. (1993). Molecular cloning and duplication of the nematode sex-determining gene *tra-1*. Genetics *133*, 543-560.

Hope, I. A. (1994). PES-1 is expressed during early embryogenesis in *Caenorhabditis elegans* and has homology to the fork head family of transcription factors. Development *120*, 505-514.

Hosono, R., Hekimi, S., Kamiya, Y., Sassa, T., Murakami, S., Nishiwaki, K., Miwa, J., Taketo, A., and Kodaira, K. -I. (1992). The *unc-18* gene encodes a novel protein affecting the kinetics of acetylcholine metabolism in the nematode *Caenorhabditis elegans*. J. Neuro. *58*, 1517-1525.

Huang, X. -Y., and Hirsh, D. (1989). A second trans-spliced RNA leader sequence in the nematode *Caenorhabditis elegans*. PNAS (USA) *86*, 8640-8644.

Ishii, N., Wadsworth, W. G., Stern, B. D., Culotti, J. G., and Hedgecock, E. M. (1992). UNC-6, a laminin-related protein, guides cell and pioneer axon migrations in *C. elegans*. Neuron *9*, 873-881.

Krause, M., and Hirsh, D. (1987). A trans-spliced leader sequence on actin mRNA in *C. elegans*. Cell *49*, 753-761.

Lee, J., Jongeward, G. D., and Sternberg, P. W. (1994). *unc-101*, a gene required for many aspects of *Caenorhabditis elegans* development and behavior, encodes a clathrin-associated protein. Genes & Develop. *8*, 60-73.

Leung-Hagesteijn, C., Spence, A. M., Stern, B. D., Zhou, Y., Su, M. -W., Hedgecock, E. M., and Culotti, J. G. (1992). *UNC-5*, a transmembrane protein with immunoglobulin and thrombospondin type-1 domains, guides cell and pioneer axon migrations in *C. elegans*. Cell *71*, 289-299.

Lewis, J. A., Wu, C. -H., Levine, J. H., and Berg, H. (1980). Levamisole-resistant mutants of the nematode *Caenorhabditis elegans* appear to lack pharmacological acetylcholine receptors. Neuroscience *5*, 967-989.

Liao, L. W., Rosenzweig, B., and Hirsh, D. (1983). Analysis of a transposable element in *Caenorhabditis elegans*. PNAS (USA) *80*, 3585-3589.

MacLeod, A. R., Karn, J., and Brenner, S. (1981). Molecular analysis of the *unc-54* myosin heavy-chain gene of *Caenorhabditis elegans*. Nature *291*, 386-390.

McIntire, S. L., Jorgensen, E., and Horvitz, H. R. (1993). Genes required for GABA function in *Caenorhabditis elegans*. Nature *364*, 334-337.

Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. EMBO J. *10*, 3959-3970.

Mendel, J. E., Korswagen, H. C., Liu, K. S., Hadju-Cronin, Y. M., Simon, M. I., Plasterk, R. H. A., and Sternberg, P. W. (1995). Participation of the protein G_0 in multiple aspects of behavior in *C. elegans*. Science *267*, 1652-1655.

Miller, D. M., Niemeyer, C. J., and Chitkara, P. (1993). Dominant *unc-37* mutations suppress the movement defect of a homeodomain mutation in *unc-4*, a neural specificity gene in *Caenorhabditis elegans*. Genetics *135*, 741-753.

Moerman, D. G., Benian, G. M., and Waterston, R. H. (1986). Molecular cloning of the muscle gene *unc-22* in *Caenorhabditis elegans* by Tc1 transposon tagging. Proc. Natl. Acad. Sci. USA *83*, 2579-2583.

Mori, I., Moerman, D. G., and Waterston, R. H. (1990). Interstrain crosses enhance excision of Tc1 transposable elements in *Caenorhabditis elegans*. Mol. Gen. Genet. *220*, 251-255.

Nonet, M. L., Grundahl, K., Meyer, B. J., and Rand, J. B. (1993). Synaptic function is impaired but not eliminated in *C. elegans* mutants lacking synaptotagmin. Cell *73*, 1291-1305.

Ogura, K., Wicky, C., Magnenat, L., Tobler, H., Mori, I., Müller, F., and Ohshima, Y. (1994). *Caenorhabditis elegans unc-51* gene required for axonal elongation encodes a novel serine/threonine kinase. Genes & Dev. *8*, 2389-2400.

Pilgrim, D. (1993). The genetic and RFLP characterization of the left end of linkage group *III* in *Caenorhabditis elegans*. Genome *36*, 712-724.

Riddle, D. L., Swanson, M. M., and Albert, P. S. (1981). Interacting genes in nematode dauer larva formation. Nature *290*, 668-671.

Rogalski, T. M., Williams, B. D., Mullen, G. P., and Moerman, D. G. (1993). Products of the *unc-52* gene in *Caenorhabditis elegans* are homologous to the core protein of the mammalian basement membrane heparan sulfate proteoglycan. Genes & Dev. *7*, 1471-1484.

Rose, A. M., Harris, L. J., Mawji, N. R., and Morris, W. J. (1985). Tc1(Hin): a form of the transposable element Tc1 in *Caenorhabditis elegans*. Can. J. Biochem. Cell Biol. *63*, 752-756.

Rosenzweig, B., Liao, L. W., and Hirsh, D. (1983). Sequence of the *C. elegans* transposable element Tc1. Nucleic Acids Research *11*, 4201-4209.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning, a laboratory manual*, 2nd. edition. Cold Spring Harbor Laboratory Press, NY.

Schnabel, H., and Schnabel, R. (1990). An organ-specific differentiation gene, *pha-1*, from *Caenorhabditis elegans*. Science *250*, 686-688.

Ségalat, L., Elkes, D. A., and Kaplan, J. M. (1995). Modulation of serotonin-controlled behaviors by G_0 in *Caenorhabditis elegans*. Science *267*, 1648-1651.

Siddiqui, S. S., and Culotti, J. G. (1991). Examination of neurons in wild-type and mutants of *Caenorhabditis elegans* using antibodies to horseradish peroxidase. J. Neurogenet. *7*, 193-211.

Sulston, J., Albertson, D. G., and Thomson, J. N. (1980). The *Caenorhabditis elegans* Male: Postembryonic Development of Nongondal Structures. Dev. Biol. *8*, 542-576.

Thomas, J. H., Birnby, D. A., and Vowels, J. J. (1993). Evidence for Parallel Processing of Sensory Information Controlling Dauer Formation in *Caenorhabditis elegans*. Genetics *134*, 1105-1117.

White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1986). The structure of the nervous system of *Caenorhabditis elegans*. Philos. Trans. R. Soc. Lond. B Biol. Sci. *314*, 1-340.

Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., *et al*. (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome *III* of *C. elegans*. Nature *368*, 32-38.

Wood, W. B. (Editor), (1988). *The Nematode Caenorhabditis elegans*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

2.6 Addendum

The data in this addendum constitute experiments not fully described in the body of the chapter, and related results that have been obtained since publication. In particular, the following are described: the mutagenesis scheme used to obtain new *unc-119* alleles; electron micrographs of wild-type and mutant muscle; placement of *unc-119* in the dauer formation genetic epistatic pathway; errors in the nucleotide sequence of the *unc-119* gene; a stem-loop strucure in the *unc-119* 5'UTR; the altered mRNA predicted by the lesion in *unc-119*(*ed9*); the earliest expression of the *unc-119*::lacZ transgene; the cloning and sequencing of the *eDf2* breakpoint, and the identification of UNC-119 sequence homologues.

2.7 Screen for new *unc-119* mutations

The basic genetic noncomplementation screen used to obtain the *unc-119* alleles *ed3*, *ed4* and *ed9* was described in §2.2.2. A flow diagram of this strategy has been included here for clarity (Figure 2.10). Briefly, animals heterozygous for the deficiency *tDf2* and a *dpy-18* marker chromosome were mutagenized and allowed to self-fertilize. The self-progeny normally consist of 1/4 arrested embryos (putative homozygoud deficiency animals), 1/2 wild types, and 1/4 Dpy animals. Any new mutation in *unc-119*, which must occur on a *dpy-18* chromosome, should appear opposite the *tDf2* chromosome 50% of the time. Such hemizygotes will show the Unc-119 phenotype and can be picked for propagation; those breeding true are kept for further analysis.

To remove unlinked mutations free of the *dpy-18* and *tDf2* lesions, the strain is backcrossed to wild type (N2). Unc non Dpy animals are obtained and self-fertilized once more to yield *unc-119* homozygotes for further study.

2.8 Electron micrographs of wild-type and *unc-119* muscle

Although the locomotory defect seen in *unc-119* animals (ventral coiling) is of a type characteristic of neural Unc mutations (Hodgkin, 1997), direct examination of sarcomere ultrastructure was used to rule out muscle defects. For example, mutations in *unc-54*, the gene encoding the major myosin isoform of *C. elegans*, result in sarcomeres with only 30% the wild-type number of thick filaments (MacLeod *et al.* 1977).

Representative micrographs of mid-body transverse sections of wild type (N2) and *unc-119(e2498)* animals are shown in Figures 2.11 and 2.12, respectively. Several structures are indicated, deduced by comparison with published micrographs (Waterston, 1988). The wild-type thin section was inferior to that of *unc-119*, resulting in an image of lower quality. The body muscles in *unc-119* animals appear to be completely normal. Attachment to the body wall, the location and placement of dense bodies, and the array of actin and myosin filaments are similar to those seen in the wild-type section, and in other published micrographs of wild-type muscle (Waterston, 1988). This result is consistent with a neural basis for the *unc-119* phenotype, although we cannot completely rule out very subtle effects on muscle.

2.9 Placement of *unc-119* in the dauer formation pathway

Mutants affecting dauer larva formation exist in two classes: those that form dauer larvae consitutively (Daf-c), and those that are unable to form dauer larvae (Daf-d; Golden and Riddle, 1984). Most of the Daf genes have been placed in a genetic epistatic pathway based on the phenotype of Daf-c; Daf-d double-mutant combinations (Riddle and Albert, 1997). As *unc-119* mutants are Daf-d, double mutants were made with *unc-119*(*e2498*) and a second Daf-c gene, either *daf-7*(*e1372ts*) or *daf-11*(*m47ts*) (see §2.2.4).

Both *daf-7*(*e1372ts*) and *daf-11*(*m47ts*) animals form dauer larvae constitutively at 25°C (Golden and Riddle, 1984). The double mutant *unc-119*; *daf-7* formed SDS-resistant dauer larvae at 25°C, while *unc-119* ; *daf-11* animals did not. This was consistent with the placement of *unc-119* in the dauer formation pathway at a position downstream of *daf-11*, but upstream of *daf-7* (Figure 2.13). This position coincides with mutations that affect the structure of chemosensory neurons, or Dyf (dye-filling) mutants (Perkins *et al.*, 1986; Starich *et al.*, 1995). As anticipated by this result, it was later confirmed that *unc-119* mutants have a dye-filling defect (see Chapter 4).

2.10 Onset of *unc-119*::lacZ expression

A description of the construction and expression of the *unc-119*::lacZ reporter transgene was given in §2.2.6 and §2.3.5, respectively. The earliest expression of the *unc-119*::lacZ transgene has implications for the possible role of *unc-119* in neural development, as we might expect that the expression of the gene anticipates the requirement for UNC-119 function. Also, as neural *unc-119*::lacZ expression is observed in adults, we can predict that embryonic expression might occur in neural cell precursors (Sulston *et al.*, 1983).

An attempt to determine the earliest time of expression was performed as follows. Slides containing animals fixed and stained for DNA and β-galactosidase activity (see §2.2.7) were scanned for young embryos as judged by counting of DAPI-staining nuclei. Several embryos (n=5), with approximately 60-80 nuclei, showed blue staining in four adjacent cells. The number of nuclei suggest that these embryos are approximately 150 minutes post-fertilization (Deppe, 1978). Slightly older embryos appeared to show staining in eight cells bilaterally opposite each other as two sets of four. In older embryos, the number of expressing cells quickly increased. Since these preparations did not preserve cellular ultrastructure, this made identification of individual cells difficult.

Two of the youngest embryos were digitally photographed in four focal planes by Nomarski optics. By overlaying the outlines of nuclei from each layer, a diagram of all of the nuclei of one of these embryos was constructed (Figure 2.14). By comparison with a diagram of nuclei in a 100-minute embryo (Sulston *et al.*, 1983), the position of the *unc-119*::lacZ expressing cells suggests that they are descendants of the AB blastomere, in particular, the granddaughters of ABpr. This is not surprising, as most of the ABpr descendants differentiate into neurons (Sulston, 1983), and *unc-119*::lacZ expression occurs in a large number of neurons in the adult (Figure 2.9). Expression in the granddaughters of ABpr alone is not sufficient to account for all the neurons expressing *unc-119*::lacZ in the adult, however. For example, the ALM and CAN neurons, which express *unc-119*::GFP (Figure 4.15), descend from ABarpp and ABalap, respectively (Sulston, 1983). No further experiments were performed to confirm the identify of these cells. As mutation in *unc-119* affects neurite outgrowth (see Chapter 4), it is unclear what role UNC-119 might be playing, if any, in pre-differentiation embryos.

2.11 Inaccuracies in the *unc-119* sequence

Subsequent to the publication of this chapter, the genome sequencing project completed the sequence of the *unc-119*-containing cosmid M142. When the sequence obtained in this work is compared to that obtained by the sequencing project, several discrepancies are seen in the form of missing bases, inserted bases and substitutions (Table 2.1). The original autoradiographs from which these sequences were read were re-examined, and most of the discrepancies were found to fall into one of three categories: compression artifacts, 4-lane termination artifacts, and misreading. No omissions or substitutions greater than four bases were found. All of these errors (with the possible exception of one) were mine. However, the regions that correspond to the cDNAs were completely confirmed by the M142 sequence.

Mistakes were a direct result of the approach taken to sequence the gene, which primarily consisted of using universal primers on restriction fragment subclones with single sequence reaction runs. The accuracy of the coding sequence resulted from the redundancy obtained by sequencing the cDNAs in parallel, and the reduced likelihood of repeated bases (and hence, potential sites for sequence errors) in coding sequences.

2.12 Stem-loop structure in the *unc-119* 5'UTR

The *unc-119* cDNA contains an apparently unused start codon (§2.3.4). This start codon was found to reside in a sequence with the potential to form a strong "stem-loop" secondary structure, prompting speculation of a possible role in regulation. The structure starts 50 bases downstream of the transcription start site, has a stem of eight perfect base pair matches and a loop of seven nucleotides, and a predicted ΔG of -16.0 kcal/mol (Figure 2.15). There does not appear to be a role for this structure, as an unused ATG and stem-loop do not occur in the homologous *C. briggsae* sequence (Chapter 3). Furthermore, *unc-119*::GFP fusions that exclude this portion of the message still drive expression in the nervous system (pUGF9; data not shown). It may simply be that this structure, formed in the mature mRNA, prevents ribosomal access to the unused AUG.

Another such region, an 8-bp stem and 10-base loop, occurs 615 bases downstream of +1, within the UNC-119 coding region.

2.13 Altered *unc-119* mRNA in the *ed9* allele

The lesion in the *ed9* allele of *unc-119* altered the conserved 5'-AG-3' to 5'-AA-3' at the splice acceptor at the right-hand junction preceding exon IV (§2.3.4). The next base, found at the beginning of exon IV, is a guanine, which is predicted to produce a candidate 5'-AAG-3' acceptor sequence (Krause, 1995). The mature mRNA would lack the G normally found in exon IV, effectively producing a frame shift mutation.

To confirm the production of such a frameshifted message, a portion of the *ed9* cDNA was amplified by RT-PCR using the primers MMA3 and MMA1 (see §4.2). A fragment of the PCR product was cloned and sequenced. As predicted, the cDNA was missing a single base, resulting in a (-1) translational frameshift in the UNC-119 coding region. The predicted polypeptide produced in *ed9* and the other *unc-119* alleles is shown in Figure 2.16. The shortened UNC-119 ORF in the *ed9* allele is consistent with its apparent null phenotype (§2.3).

2.14 The molecular identification of the *eDf2* breakpoint

2.14.1 Background

The CB1517 strain, which carries *eDf2* and *eDp6*, was recovered after mutagenesis with acetaldehyde (Hodgkin, 1980). This resulted in the separation of LG*III* into two independently segregating pieces, *eDf2*, which carries about 2/3 of the chromosome, and *eDp6*, which carries the rightmost 1/3, based on preliminary genetic analysis. Subsequent mutagenesis studies with gamma irradiation suggested that *eDp6* is actually a circular chromosome (C. Hunter and W. Wood, pers. comm.).

Until recently, loci in the region could be mapped to either the left or right of the breakpoint by the ability of a mutation on LG*III* to be complemented by either *eDf2* or *eDp6*. The CB1517 strain is strongly Unc, and as mentioned earlier, it was this similar phenotype of CB1517 and *unc-119*(*e2498*) which led to initial speculation that this was the gene disrupted by the aberrations.

The pDP#MM008 fragment, which contains the majority of the *unc-119* gene, demonstrated very little hybridization to genomic DNA of CB1517 (Figure 2.6), suggesting that the *eDf2* breakpoint lay within the pDP#MM008 region. A smaller region was identified by using portions of pDP#MM008 as probes on Southern blots of CB1517 DNA (data not shown). This localized the breakpoint to within a ~1.2 kb *Eco*RI fragment, the same region which contains the transcription initiation site of *unc-119* (shown in Figure 2.16). Although this precludes expression of any part of the UNC-119 ORF, it was of interest to determine the nature of the breakpoint sequences, as the potential existed for a novel gene fusion of *unc-119* with the sequences added to the right of the*eDf2* breakpoint.

2.14.2 Cloning of the eDf2 breakpoint

Further Southern hybridization analysis indicated that the breakpoint was contained in an *Sst*I-*Pst*I fragment of size 5.5 kb in CB1517 (data not shown). An *Sst*I-*Pst*I plasmid minilibrary was constructed from CB1517 DNA, and probed with the leftmost *Hin*DIII-*Eco*RI fragment of pDP#MM008. One clone from the minilibrary which hybridized was called pDP#MM071. A restriction map of this clone confirmed the earlier Southern hybridization data and placed the *eDf2* breakpoint within the same region as previously determined. Several subclones of pDP#MM071 were obtained and sequenced. A ~1.4 kb *Eco*RV-*Bam*HI subclone, pDP#MM076, allowed the precise identification of the *eDf2* breakpoint (Figures 2.17 and 2.18).

A preliminary database search with the sequences to the right of the breakpoint revealed no significant similarities to known nucleotide sequences, although a putative splice acceptor sequence could be identified, upstream of a short coding region with homology to zinc finger transcription factors (data not shown). The putative exon is indicated in Figure 2.17. To determine the source of the DNA added to the right of the *eDf2* breakpoint, a 0.9-kb *Cla*I-*Xba*I fragment of pDP#MM071, to the right of the breakpoint, was used to probe a *C. elegans* YAC grid (Coulson *et al.*, 1995). Cross-hybridization was obtained with two YAC clones, Y76A2 and Y30E6. Five cosmids mapping to these YACs were obtained from the *C. elegans* physical mapping project; one of them, C30E9, contained DNA from the right-hand half of pDP#MM071. YACs Y76A2 and Y30E6 are from a contiguous set of overlapping clones localized to the right end of linkage group *III*, as indicated by the *C. elegans* physical map information in ACeDB/Macace (Eeckman and Durbin, 1995). The conclusion is that *eDf2* is a large interstitial deletion, and not a terminal deletion.

Subsequent to the localization of the *eDf2* breakpoint to cosmid C30E9, the sequence of the overlapping cosmid T05D4 became available from the *C. elegans* genome sequencing consortium. The sequence to the right of the breakpoint was identical to that obtained from pDP#MM076 except for the presence of an additional base pair, attributable to a sequence error on my part (data not shown). A comparison of wild-type sequences around the breakpoint revealed no strong similarities that might suggest a homologous intrachromosomal recombination origin for *eDf2* (Figure 2.18).

2.14.3 LGIIIR is duplicated in CB1517

Unexpected results were obtained upon characterization of the right half of pDP#MM071. When used as a probe against Southern blots of *Sst*I-digested CB1517 DNA, the polymorphic fusion band showed cross-hybridization, but a band of similar size as N2 was also found, suggesting that the DNA appended to the *eDf2* breakpoint was still intact in CB1517 (Figure 2.19). Contamination of the CB1517 DNA preparation can be ruled out, as the probe should have also revealed the wild-type band corresponding to the *unc-119* gene. Furthermore, the presence of both polymorphic and wild-type bands was subsequently demonstrated in CB1517 by hybridization to Southern blots of digestions with the enzymes *Eco*RV, *Eco*RI, *Pst*I and *Hin*DIII (data not shown).

If the DNA that was joined to the right of the *eDf2* breakpoint is also present without aberration, it follows that the potential disruption of a putative second gene associated with the right portion of *eDf2* might not result in a phenotype. This is consistent with the ability of *unc-119* rescuing clones to restore a wild-type phenotype to *eDf2; eDp6* animals (see Figure 2.7), and suggests that the potential expression of the fusion sequences under control of the *unc-119* promoter does not cause any obvious phenotype.

2.14.4 Models for the origin of eDf2 and eDp6

The simplest model consistent with these data are intrachromosomal two-break model for the creation of *eDf2*, and the subsequent circularization of the intervening portion to form a circular chromsome, *eDp6* (Figure 2.20A). This model is contradicted, however, by the finding that the wild-type sequences, which correspond to the novel fusion in *eDf2*, are found in CB1517. The possibility that at least part of the far right end of LG*III* is also present on *eDp6* is supported by one piece of genetic evidence: The mutation *dyf-2*(*m160*), which maps to the far right end of LG*III*, is complemented by both *eDf2* and *eDp6* (Starich *et al.*, 1995).

Given the physical and genetic data which imply that *eDf2* and *eDp6* partially overlap, any model which explains the origin of these aberrations must involve at least two steps. This is not unprecedented, as complex rearrangements are typical in the generation of free duplications in *C. elegans* (McKim and Rose, 1994). A possible mechanism is that the original LG*III* suffered a break in the *unc-119* region as a result of the mutagen acetaldehyde; the right-hand piece circularized to form *eDp6*, while the left-hand piece underwent non-homologous recombination towards the far right end of the other LG*III* homologue (Figure 2.20B). The other chromosome would be expected to be in close physical proximity due to meiotic pairing.

Further attempts were made to clone polymorphic DNA fragments from CB1517. The arrangement of the putative *eDp6* breakpoint fragments showed that this region contained numerous rearragements associated with the adjacent sequences, which preclude construction of a simple model for the creation of *eDp6*. Hence, the precise *eDp6* breakpoint, which presumably contains the site of circularization, has not been identified.

2.15 Homologues of UNC-119

At the time of publication of the sequence of UNC-119, no proteins were known with significant similarity to UNC-119, except for the predicted gene C27H5.1 (see §2.3.4 and §3.9). Shortly after, two new homologues became known, HRG4 (for human retinal gene 4) and its homologue in rat, RRG4 (Higashide *et al.*, 1996). The degree of conservation with these homologues, and their functional relatedness, will be discussed in Chapter 4.

position	my sequence	M142	type of error	attributed cause
161	CAACGT	CAAACGT	missing base	poor gel resolution
376	GGAAACG	GGAAAAACG	missing bases	4-lane stop in gel
720	GAACGATTG	GAAGGCAATTG	substitution/ insertion	compression in gel
1011	TCCTCTCTTCT	TCCCTCTCTTTCT	missing bases	poor gel resolution
1037	TCTCTTTCT	TCTTTCT	extra bases	misread repeated pattern
1152	GCCTTTTTC	GCCCTTTTTC	missing base	multiple bases
2361	ACGAAGAAAT	ACGTTGAAAT	substitution	misreading of sequence
2727	TGGTTGGC	TGTTCGGC	substitution	poor gel resolution
2913	GCTTTAAGGC	GCAAATTGGC	substitution	misreading of sequence
2933	AACCCAAAGT	AATTTAAAGT	substitution	misreading of sequence
3090	ATTTTGCC	ATTTGCC	extra base	4-lane stop
3098	ACCAAC	ACGGCAAC	missing bases	4-lane stop
3132	AGCTGC	AGCGTGC	missing base	poor gel resolution
3182	TTGTGAAAT	TTGAAAT	extra bases	may not be an error?
4132	TCTCGGAA	TCTGGAA	extra base	error in 1° Y52D3 read
4903	AATTTACCG	AATTCCG	extra bases	poor gel resolution
4951	TAAACTTC	TAAAACTTC	missing base	poor gel resolution
4964	CAAATTC	CAAAATC	substitution	poor gel resolution
4980	AGTTCAG	AGTCAG	extra base	poor gel resolution
4996	CTTGACCG	CTTTGCCG	missing base/ extra base	poor gel resolution
5007 - ~5060	GTTTCA. . . . . .GCATG	GTTTTCA. . . . . .ATTCC	multiple errors	poor gel resolution

Table 2.1 Summary of sequencing errors in originally published *unc-119* sequence (Chapter 2). The 'position' of each error is given as the distance, in base pairs, from the *Hin*DIII site (i.e. in plasmid pDP#MM016) to the corresponding bases in the *C. elegans* genome sequencing project sequence (from cosmid M142).

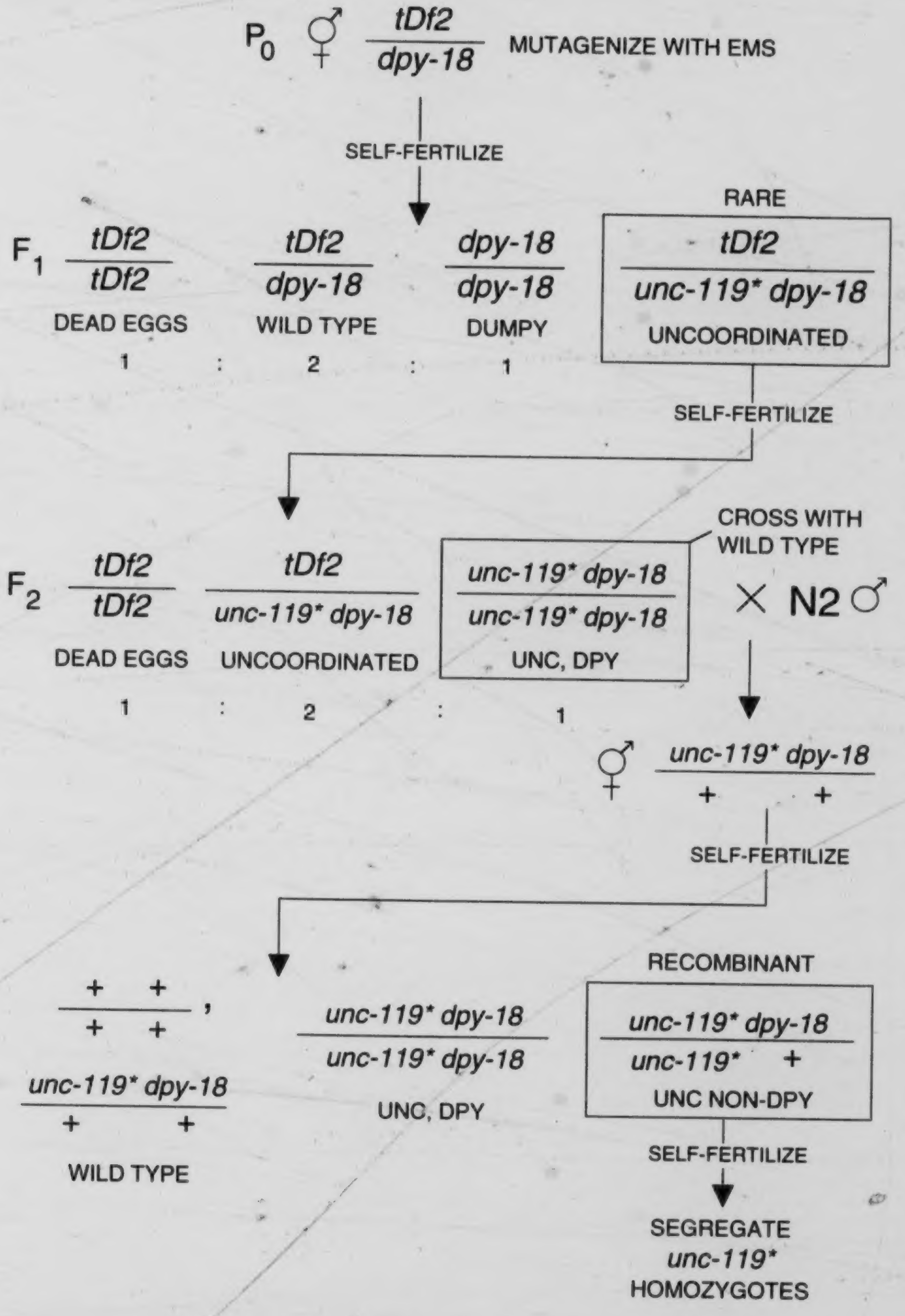


Figure 2.10 Mutagenesis scheme for generating new alleles of *unc-119* based on noncomplementation with the deficiency *tDf2*. This deficiency is homozygous lethal, and removes *unc-119*, but not *dpy-18* (see text).

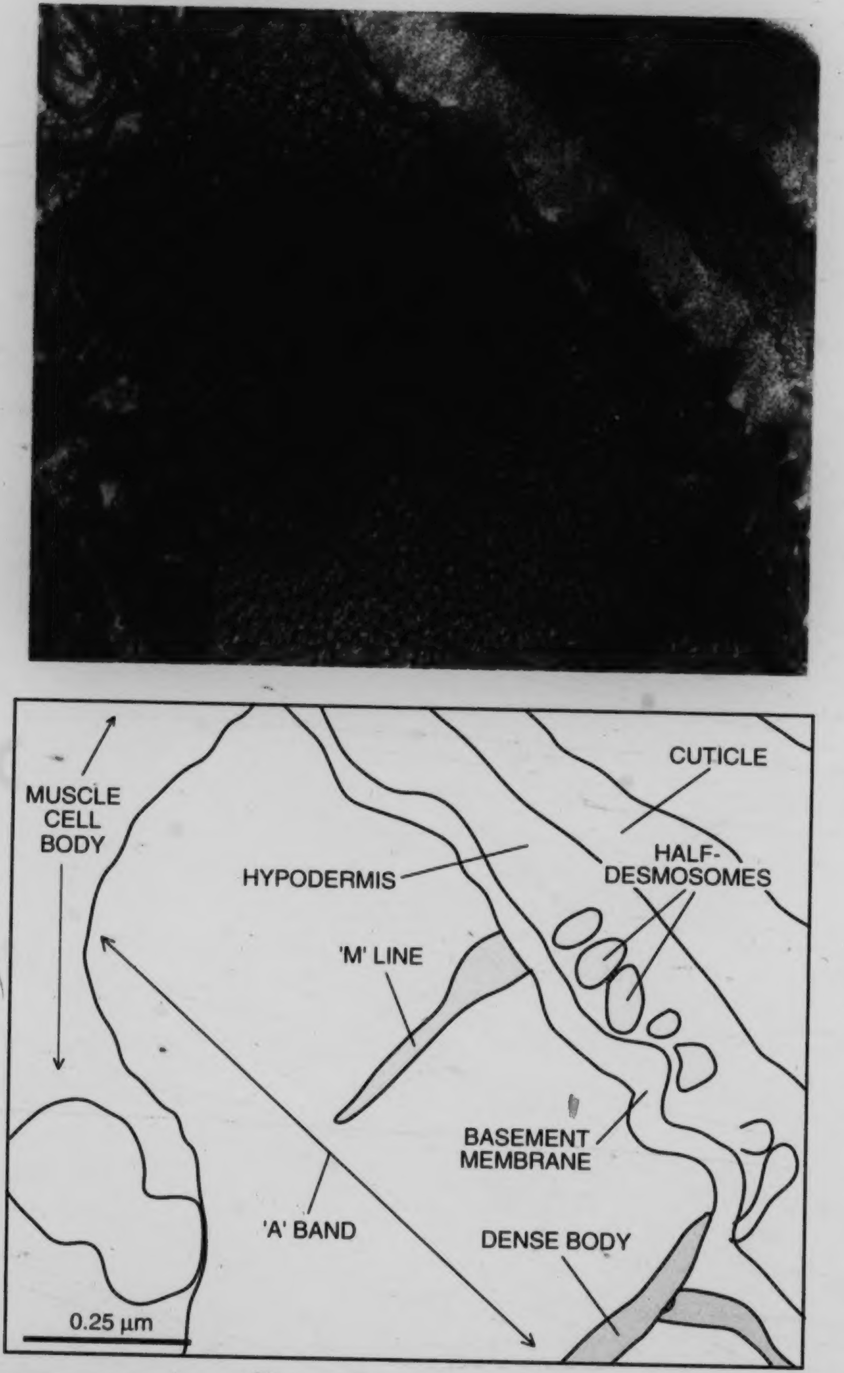


Figure 2.11 Electron micrograph of wild-type (N2) muscle sarcomere (top). A line representation of the micrograph is shown below, to indicate particular structures.

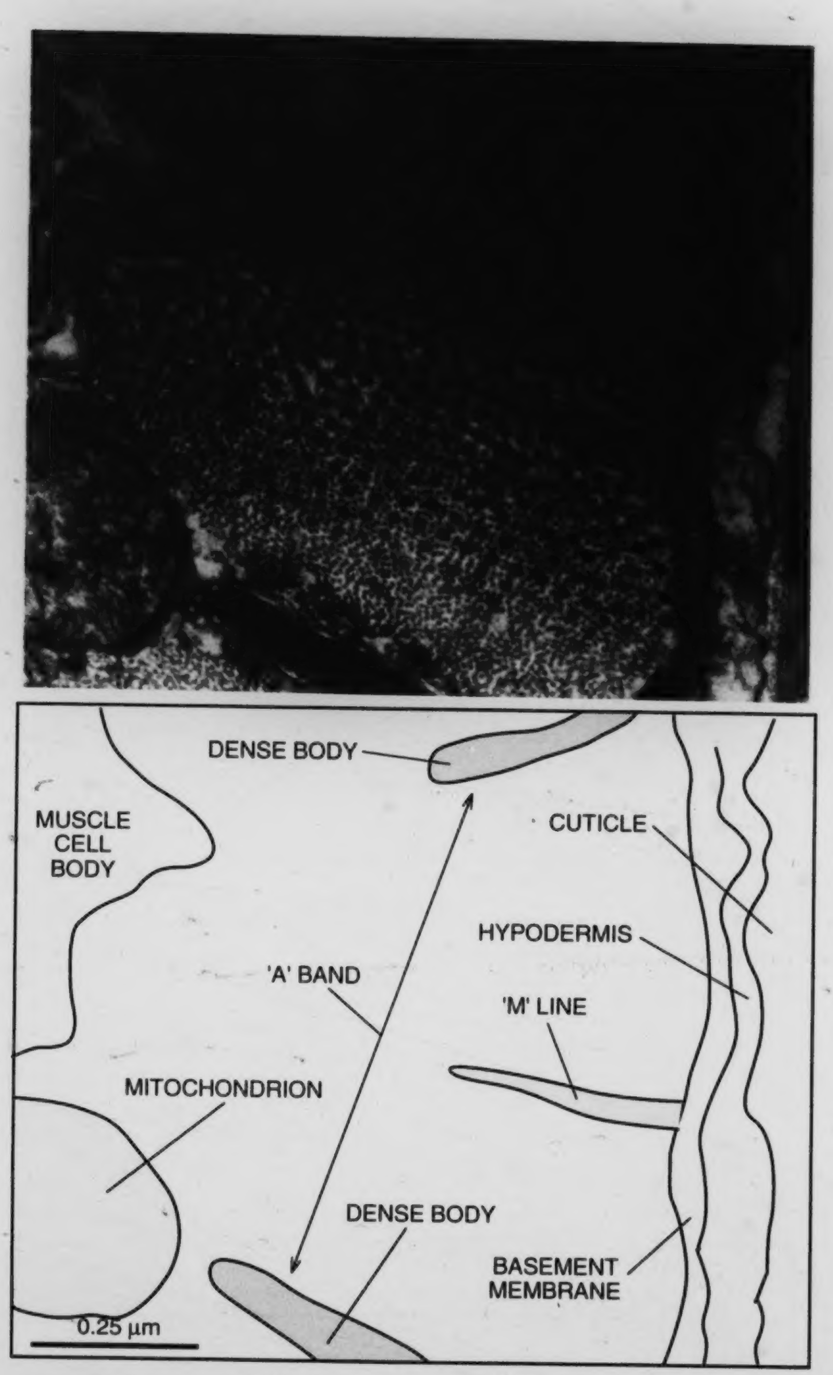


Figure 2.12 Electron micrograph of *unc-119*(*e2498*) muscle sarcomere (top). A line representation of the micrograph is shown below, to indicate particular structures.

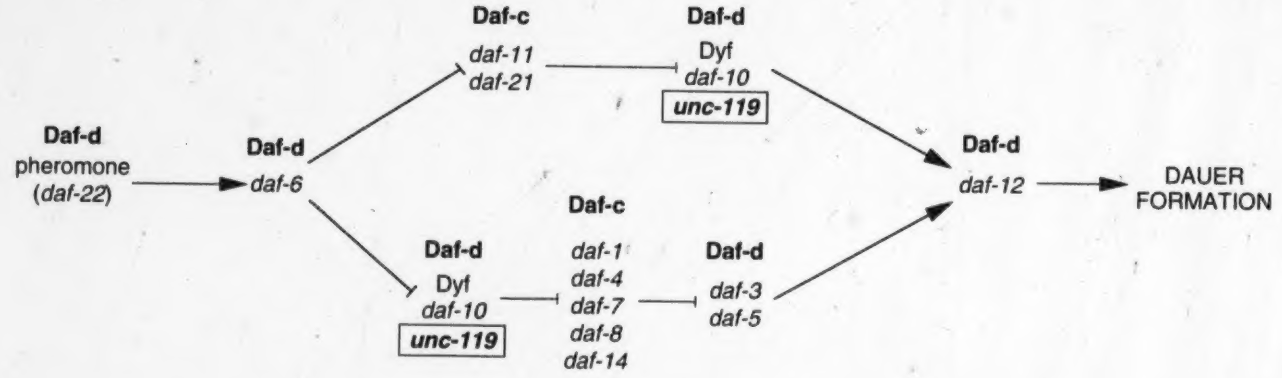


Figure 2.13 Placement of *unc-119* in the dauer formation genetic epistasis pathway. The type of dauer defect caused by loss of function mutations is indicated (Daf-d, dauer formation defective; Daf-c, dauer constitutive). Dyf mutants have abnormal dye filling of chemosensory neurons due to defective cilium structure (Perkins *et al.*, 1986; Starich *et al.*, 1995). The pathway shown, which leaves out genes contributing to adult longevity, was adapted from Riddle and Albert (1997).

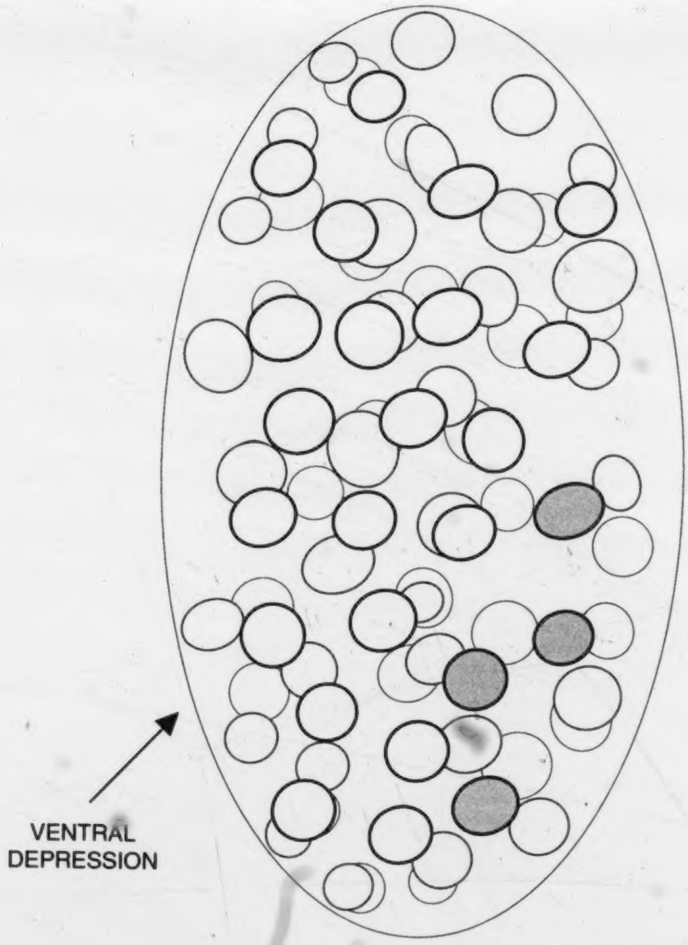


Figure 2.14 Diagram of nuclei in a wild-type embryo carrying pRF4 (rol-6D) and the *unc-119*::lacZ transgene. The embryo was deduced to be approximately 150 minutes post-fertilization (see text). Expression is observed in four unidentified nuclei (shaded). Nuclei from a similar focal plane are drawn with circles of similar thickness. The diagram was made by merging four camera lucida drawings from four sequential focal planes. The focal plane showing the lacZ expression has been arbitrarily chosen as the 'top' layer in the merged figure. Ventral is to the left, and anterior is at the top.

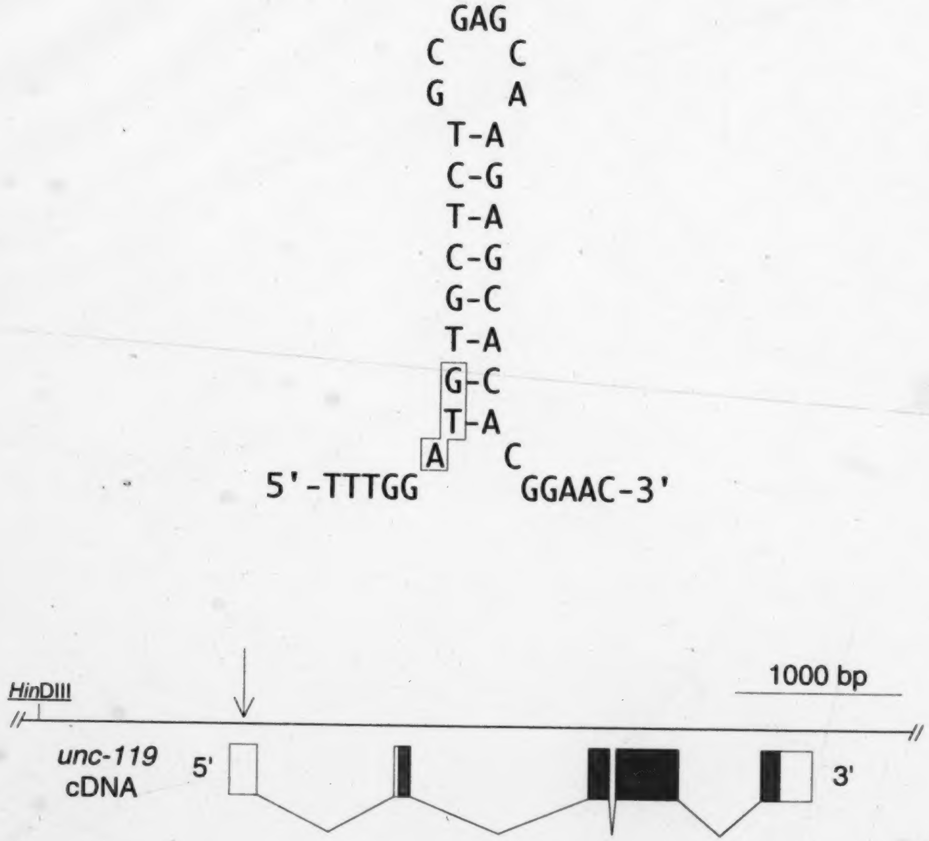


Figure 2.15 Stem-loop structure present at the 5' noncoding sequence of *unc-119* (shown by an arrow). The unused start codon is outlined by a box. The secondary structure has a predicted ΔG of -16.0 kcal/mol as determined by a public Internet server program (www.ibc.wustl.edu/~zuker), using the theoretical determinants of Turner *et al.* (1988).

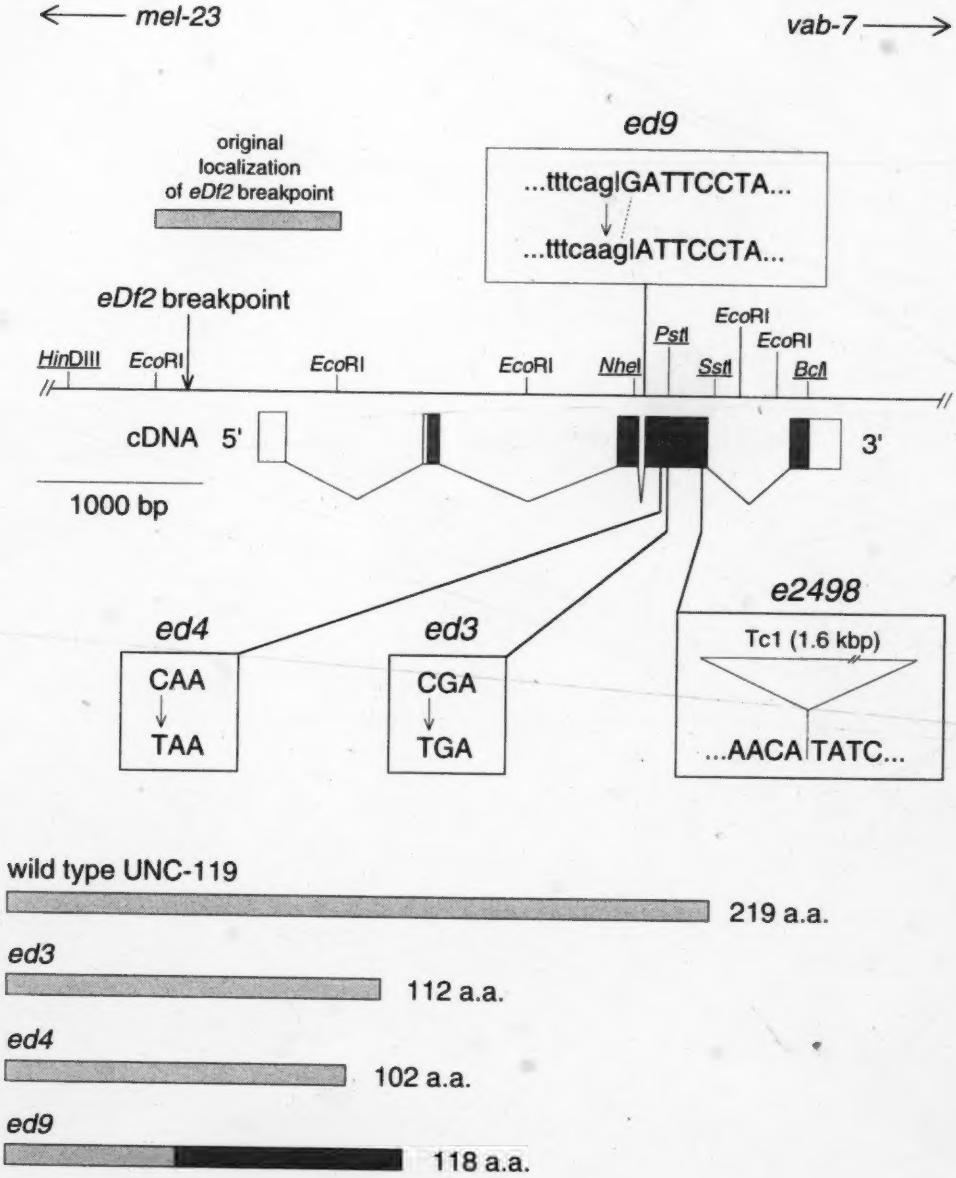


Figure 2.16 Sequence alterations in *unc-119* mutants, and predicted polypeptide sizes for the point mutants *ed3*, *ed4* and *ed9*. Altered amino acids in the predicted *ed9* gene product, shown by darker shading, result from translation of out-of-frame codons. The location of the *eDf2* breakpoint, which is upstream of the *unc-119* message, is indicated.

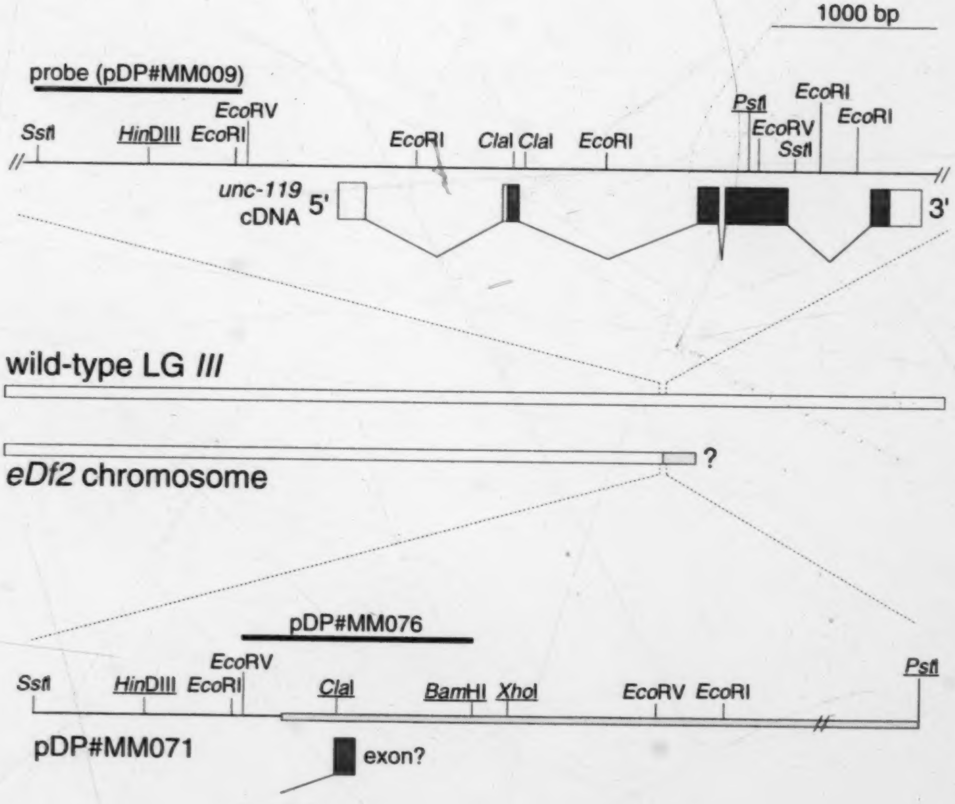
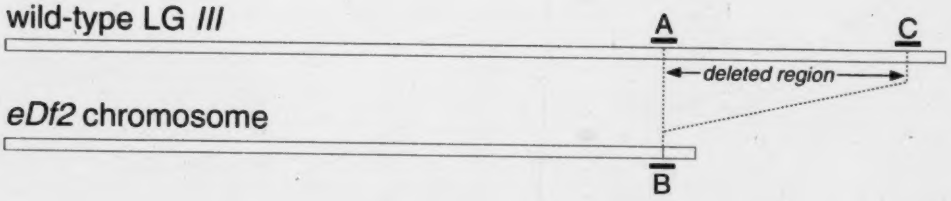


Figure 2.17 Restriction maps of the *eDf2* breakpoint-containing clone pDP#MM071, and the corresponding wild-type *unc-119* region. Unique sites are underlined. The probe used to obtain the *eDf2 SstI-PstI* clone is indicated above the *unc-119* restriction map, and the pDP#MM076 subclone containing the breakpoint is shown above the pDP#MM071 map. The sequences added to the right of the *eDf2* breakpoint are shown as shaded; the true size of the added portion in the *eDf2* chromosome is not known. All DNA is shown with the same left-right orientation as the *C. elegans* physical map.



A accggcaatttggcgaaatttgccggaaggcaattgccgcccacccctgt

B accggcaatttggcgaaatttgccgTTAAGTGACAAGAGTGGGGACACAT

C CGACACCCCTTATATATGTGTAGGATTAAGTGACAAGAGTGGGGACACAT

Figure 2.18 Sequence segments near the two *eDf2* breakpoints (A and C), and the novel join occurring in the *eDf2* chromosome (B). Upper and lower case base sequences are used to differentiate the two wild-type segments. The sequences do not reveal any strong homologies that would suggest a recombination model for the creation of *eDf2*. The distance from the right end of *III* to sequence C is not precisely known. The A and C sequence segments are from the *C. elegans* genome project sequences of cosmids M142 and T05D4, respectively; the sequence from region B is from pDP#MM076.

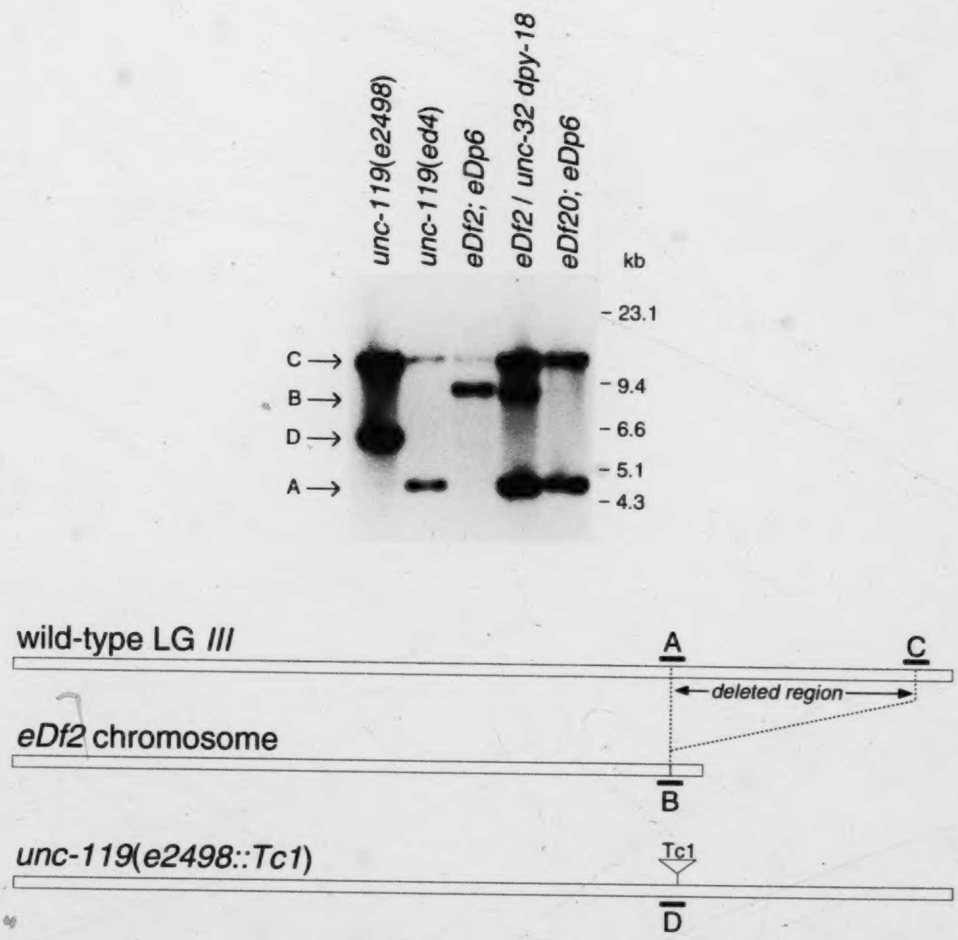


Figure 2.19 Autoradiograph of Southern hybridization of various *SstI* digests probed with the insert from pDP#MM071, a fragment spanning the *eDf2* breakpoint. The bands correspond to segments of DNA near the breakpoint region shown below the gel. Band D is shifted upwards due to the presence of the transposable element Tc1 in the *e2498* allelle. Strain *unc-119(ed4)* and the *eDf20* and *unc-32 dpy-18* chromosomes are not polymorphic in the probe region and show wild type band A. The polymorphic band B segregates with *eDf2*, but an additional wild-type band C is present in *eDf2; eDp6*. The presence of wild-type restriction fragments from the right-hand C region has been confirmed with other enzymes (data not shown). (Compare with Figure 2.4)

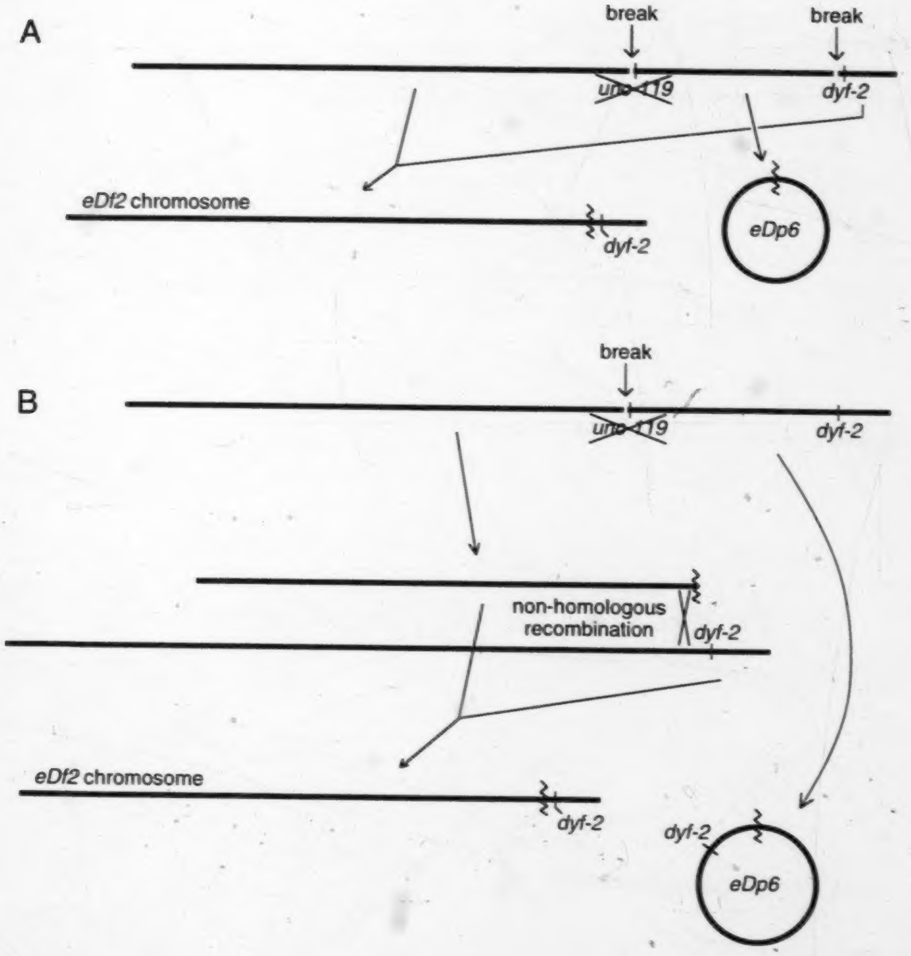


Figure 2.20 Models for the origin of *eDf2* and *eDp6*. ***A***, Two-break model. Acetaldehyde results in two breaks in LG*III*, one which destroys *unc-119*, and the other near *dyf-2*. Joining of the outer fragments creates *eDf2*, and circularization of the intervening segment creates *eDp6*. This is not consistent with the molecular and genetic data that suggest the right terminus of LG*III* is retained in *eDp6*. ***B***, One-break and recombination model. A single break occurs in *unc-119*, and this chromosome 'heals' by undergoing non-homologous recombination with the other (intact) homologue. The circularization of the DNA to the right of *unc-119* creates *eDp6*, which retains LG*IIIR* terminal sequences (including *dyf-2*).

2.16 Bibliography

Coulson, A., Huynh, C., Kozono, Y., and Shownkeen R. (1995). The Physical Map of the *Caenorhabditis elegans* Genome. in *Caenorhabditis elegans: Modern Biological Analysis of an Organism* (eds Epstein, H. F. and Shakes, D. C.) Methods Cell Biol. *48*, 533-550.

Deppe, U., Schierenberg, E., Cole, T., Krieg, C., Schmitt, D., Yoder, B., and von Ehrenstein, G. (1978). Cell lineages of the embryo of the nematode *Caenorhabditis elegans*. PNAS (USA) *75*, 376-380.

Eeckman, F. H., and Durbin, R. (1995). ACeDB and Macace. in *Caenorhabditis elegans: Modern Biological Analysis of an Organism* (eds Epstein, H. F. and Shakes, D. C.) Methods Cell Biol. *48*, 583-605.

Golden, J. W., and Riddle, D. L. (1984) A pheromone-induced developmental switch in *Caenorhabditis elegans*: Temperature-sensitive mutants reveal a wild-type temperature-dependent process. PNAS (USA) *81*, 819-823.

Higashide, T., Murakami, A., McLaren, M. J., and Inana, G. (1996). Cloning of the cDNA for a Novel Photoreceptor Protein. J. Biol. Chem. *271*, 1797-1804.

Hodgkin, J. A. (1980). More sex determination mutants of *Caenorhabditis elegans*. Genetics *96*, 649-664.

Krause, M. (1995). Transcription and Translation. in *Caenorhabditis elegans: Modern Biological Analysis of an Organism* (eds Epstein, H. F. and Shakes, D. C.) Methods Cell Biol. *48*, 483-512.

MacLeod, A. R., Waterston, R. H., Fishpool, R. M., and Brenner, S. (1977) Identification of the structural genes for a myosin heavy-chain in *Caenorhabditis elegans*. J. Mol. Biol. *114*, 133-140.

McKim, K. S., and Rose, A. M. (1994). Spontaneous duplication loss and breakage in *Caenorhabditis elegans*. Genome *37*, 595-606.

Perkins, L. A., Hedgecock, E. M., Thomson, J. N., and Culotti, J. G. (1986). Mutant sensory cilia in the nematode *Caenorhabditis elegans*. Dev. Biol. *117*, 456-487.

RIddle, D. L., and Albert, P. S. (1997). Regulation of Dauer Larva Development. in *C. elegans II*. Cold Spring Harbor Laboratory, New York.

Starich, T. A., Herman, R. K., Kari, C. K., Yeh, W. -H., Schackwitz, W. S., Schuyler, M. W., Collet, J., Thomas, J. H., and Riddle, D. L. (1995). Mutations affecting the chemosensory neurons of *Caenorhabditis elegans*. Genetics *139*, 171-188.

Sulston, J. E. (1983). Neuronal cell lineages in the nematode *Caenorhabditis elegans*. CSHSQB *48*, 443-452.

Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. Dev. Biol. *100*, 64-119.

Turner, D. H., and Sugimoto, N. (1988). RNA structure prediction. Annu. Rev. Biophys. Chem. *17*, 167-192.

Waterston, R. H. (1988). Muscle. in *The nematode* Caenorhabditis elegans. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

3 Conservation of function and expression of *unc-119* from two *Caenorhabditis* species despite divergence of non-coding DNA

3.1 Introduction

Sequence comparisons of protein and gene sequences between species are frequently used to establish conserved elements that may have importance for protein function or gene expression (e.g. Pilgrim *et al.*, 1995). In the nematode, *C. elegans*, a project to sequence the entire genome is proceeding rapidly (reviewed in Hodgkin *et al.*, 1995), and the data are analyzed concurrently for sequences predicted to be genes. In over 20 Mbp of contiguous sequence, 4000 predicted protein coding regions have been found, of which 45% show sequence similarity to known genes (Berks *et al.*, 1995). As in the "genome projects" of other organisms, there is a large proportion of predicted genes whose function and expression are unknown. It may be possible to assign a biological role to these sequences by examination of their mutant phenotypes following targeted mutagenesis (Hodgkin *et al.*, 1995; Burns *et al.*, 1994). Alternatively, comparison of conserved regions between "sibling species" can be used to identify, by conservation, those regions likely to be important for function and regulation.

Comparison of *C. elegans* genes to those of the related species *C. briggsae* has demonstrated that while coding sequences are often highly

---

A version of this chapter has been published. Maduro M and Pilgrim D (1996). Gene 183:77-85.

conserved (allowing for degeneracy), intronic and flanking sequences have completely diverged (Snutch, 1984; Prasad and Baillie, 1989). The *C. briggsae* homologue of a *C. elegans* gene can often be detected by low stringency hybridization (Zucker-Aprison and Blumenthal, 1989; Kuwabara and Shah, 1994), and in some cases, similar sequences in the 5' flanking regions have been seen, suggesting that the control mechanisms for tissue or cell-specific gene expression have been maintained (Zucker-Aprison and Blumenthal, 1989).

We have previously described a novel protein encoded by the *C. elegans unc-119* gene, which is expressed throughout the nervous system (Maduro and Pilgrim, 1995). The phenotype of *unc-119* mutants consists of defects in nervous system function, however no biochemical role for the UNC-119 protein has yet been assigned. Since all known alleles of *unc-119* are molecular nulls, and the predicted UNC-119 protein is not significantly similar to other known proteins, there are no clues as to which regions are important for function.

The homologous *unc-119* gene in *C. briggsae* was analyzed in order to identify which parts of the gene are likely to be necessary for activity or regulation. We show that the *C. elegans* gene can be functionally replaced by the *C. briggsae* counterpart, and that despite only small sequence similarities, the control regions from each species drive βgal reporter gene expression in *C. elegans* indistinguishably.

3.2 Materials and Methods

3.2.1 Southern Analysis and Library Screening

A *C.briggsae* λCharon4 genomic library (a gift from T. Snutch and D. Baillie) was screened with purified insert DNA from pDP#MM008 (Maduro and Pilgrim, 1995) labeled with ^{32}P-dCTP using the T7 Quickprime kit (Pharmacia). The library was grown on *E. coli* LE392 and lifted to Hybond N (Amersham). Hybridization was performed at 62°C for 12 hours, followed by three 10-minute rinses at 55°C in 2 × SSC containing 0.1% SDS, prior to autoradiography. Purification of lambda and manipulation of DNA were performed as described (Sambrook, Fritsch and Maniatis, 1989) using the vector pBluescript KS- (Stratagene). Restriction enzymes were obtained from GIBCO BRL, except for *Hin*dIII (Pharmacia). Fragments of the positive λ clone pDP#MMCb1 were separated on a 0.7% agarose gel prior to blotting. Hybridization with purified *unc-119* cDNA was carried out as for the library screen.

3.2.2 5'RACE Analysis and DNA Sequencing

Sequencing was performed using Sequenase 2.0 (United States Biochemical) on small-scale single stranded templates prepared as described (Sambrook, Fritsch and Maniatis, 1989), using universal primers for the pBluescript vector. One additional oligonucleotide, MMA1 (5'-AGTCGGCCTTATTGTGCATTAC-3'), originally designed for use in *C. elegans*, was stable as a sequencing primer in *C. briggsae* despite three base mismatches (underlined). The first exon was determined by RACE using a 5'RACE System kit (BRL) according to the manufacturer's instructions. Total RNA from *C. briggsae* was prepared using a standard guanidine isothiocyanate protocol (Sambrook, Fritsch and Maniatis, 1989). The primer used for first strand synthesis was MMA21 (5'-AGTCGGCTTTGTTGTGCATGAC-3') and the nested primer was MMA5 (5'-GGAGCATAGGAATTCTTGAGTGATTCC-3'). The ~275 bp product was reamplified, cloned and sequenced as for the *C. elegans*

5'RACE product (Maduro and Pilgrim, 1995). Sequences were compiled and conceptually translated using DNA Strider 1.2 (Marck, 1988). This sequence is available through GenBank (Accession No. U45326).

3.2.3 Reporter Gene Construction

Construction of pDP#MM081, predicted to encode the first 101 aa of *C. elegans* UNC-119, has been described (Maduro and Pilgrim, 1995). For construction of the *C. briggsae* fusion, two intermediate plasmids were used to generate suitable restriction sites. The plasmid UR#91 (a gift from W. Wadsworth) contains an *Eco*RI-*Hin*dIII fragment cloned into pBluescript SK+; this fragment was removed from the polylinker of pUC19 which contained a 120-bp insertion at the *Bgl*II site. UR#91 was digested with *Eco*RI and *Bgl*II (both enzymes BRL) to liberate the 120-bp insertion and part of the pUC polylinker, and allow the insertion of a *Bam*HI-*Eco*RI fragment from *C. briggsae*. This intermediate plasmid, pDP#MM070, was digested with *Hin*dIII (Pharmacia) and *Pst*I (BRL) to allow its subcloning into similarly-digested pPD22.04 (Fire *et al.*, 1990), which is predicted to join the first 130 aa of *C. briggsae* UNC-119 with βgal.

3.3 Results And Discussion

3.3.1 Identification of the C. briggsae homologue of unc-119

A genomic phage library of *C. briggsae* was screened at low stringency using a genomic fragment containing the majority of *unc-119* from *C. elegans*. One positive clone, pDP#MMCb1, was chosen for further study. The *C. elegans unc-119* cDNA cross-hybridizes to specific restriction fragments of this clone (Figure 3.1). In order to verify that this clone contained the entire *unc-119* homologue, we injected *unc-119*(*e2498*) animals with pDP#MMCb1 and the plasmid pRF4, which contains the dominant *rol-6* marker as a control for transformation (Mello *et al.*, 1991). The *e2498* allele results from a transposon insertion in *unc-119* and has a null phenotype (Maduro and Pilgrim, 1995). Among the F_1 progeny, several Rol non-Unc animals were seen, characteristic of rescue of the *unc-119* mutant phenotype, while injection of pRF4 alone did not confer rescue.

A restriction map of the *unc-119* region of *C. briggsae* was generated, and subcloned fragments were tested for their ability to complement the Unc phenotype (Figure 3.2). When the rescuing ability of pDP#MMCb1 and a plasmid subclone was assessed in detail (Figure 3.3), defects in egg laying, pharyngeal pumping and dauer forming ability were also rescued by the *C. briggsae* transgenes, similar to results obtained with *C. elegans unc-119* clones (Maduro and Pilgrim, 1995). This confirms that the cross-hybridizing region contains a functional homologue of *C. elegans unc-119*.

3.3.2 The coding regions of the homologues are conserved

The sequence of the *C. briggsae unc-119* genomic region was determined (Figure 3.4). Using the previously determined *C. elegans* sequence and the *C. elegans* consensus splice donor and acceptor sites, we determined the presumed intron/exon junctions for the *C. briggsae* gene (shown schematically in Figure 3.2). With the exception of the first exon (see next section), determination of the intron/exon boundaries was facilitated by the high degree of coding sequence conservation. An excellent alignment between the two predicted proteins can be produced if two single amino acid (aa) gaps are allowed in the *C. briggsae* ORF. The remaining codons can be aligned with the *C. elegans* ORF and show 90% identity over 217 aa (Figure 3.5). Unfortunately, this high degree of similarity does not allow identification of specific amino acids important for function. In *C. elegans*, we have shown that the codons specified by Exon II (the first 20 aa) are partially dispensable for function, since genomic clones containing only the last three exons (Exons III-V) are sufficient for partial phenotypic rescue (Maduro and Pilgrim, 1995). However, the amino acids encoded by Exon II from the two strains are still well conserved (90% identical), suggesting that this part of the polypeptide probably has a function.

Figure 3.5 also shows an alignment of the conceptual translation product of *unc-119* from both *C. elegans* and *C. briggsae*, as well as C27H5.1, a predicted ORF identified by the *C. elegans* genome sequencing project, and noticed previously as a potential *C. elegans* homologue of *unc-119* (Maduro and Pilgrim, 1995). When compared to both the *C. briggsae* and *C. elegans* homologues, 25% identity and 40% similarity are seen. The relative positions of the introns in the coding sequence is not conserved at all with C27H5.1 (data not shown).

Since there do not appear to be any regions more conserved than others between *C. briggsae* and *C. elegans unc-119*, either there has been insufficient time for evolutionary change (unlikely, given the divergence of the intron sequences), or there is strong evolutionary pressure on the majority of the protein sequence. The latter is consistent with the estimated 20-50 Myr (million years) of evolution between the two species (Heschl and Baillie, 1990; Lee *et al.*, 1992; Kennedy *et al.*, 1994), and results of similar studies with other homologues. The 90% identity between *C. elegans* and *C. briggsae unc-119* is similar to that seen between homologues of a ubiquitin-like protein (Jones and Candido, 1993), but much higher than the 60-70% identity seen between the functional homologues of HLH-1 (Krause *et al.*, 1994) or TRA-2 (Kuwabara and Shah, 1994). It was not surprising, therefore, to find that the *C. briggsae unc-*

119 gene could complement mutants of *C. elegans*, especially given that such interspecies rescue has been seen with the more dissimilar HLH-1 homologues (Krause *et al.*, 1994).

3.3.3 Intron and flanking sequences are divergent

The location and size of the noncoding first exon of *unc-119* in *C. briggsae* was determined by 5' RACE (rapid amplification of cDNA ends; Frohman et al, 1988). This exon is smaller than that of the *C. elegans* gene (84 bp compared to 128 bp) and bears no sequence similarity except for the occurrence of T-rich sequences (Figure 3.4). While there is an apparently unused start codon in the *C. elegans* first exon (Maduro and Pilgrim, 1995), there is no counterpart in *C. briggsae*, suggesting it is not required for expression.

There is a polyadenylation signal (5'-AATAAA-3') 206 bp after the predicted translation stop codon. From comparisons made of many *C. elegans* cDNA clones (Krause, 1995), the end of the mRNA is probably 13 bases downstream. This is comparable to the 3' end of the *C. elegans* message, in which the polyadenylation signal occurs 158 base pairs after the stop codon (Maduro and Pilgrim, 1995). There is no apparent conservation between the translation stop and the end of the message, suggesting that unlike some *C. elegans* genes (e.g. *tra-2*; Goodwin *et al.*, 1993), the expression of *unc-119* probably does not involve regulation by mRNA stability through the interaction of trans-acting factors with the 3'UTR.

The intron sizes are quite different between the two species (shown schematically in Figure 3.2). The intron between exons I and II is 890 bp long in *C. elegans*, but only 275 bp in *C. briggsae*. The largest difference is in the next intron, 1070 bp in *C. elegans* but only 83 bp in *C. briggsae*. The other introns are closer in length, with *C. elegans/C.briggsae* sizes of 161/43 and 488/51 bp. This pattern of smaller introns in *C. briggsae* has been seen in many genes, including *ges-1* (Kennedy *et al.*, 1993), where the average *C. elegans* intron is eight times as large.

3.3.4 Putative promoter elements are conserved in the 5' flanking region

Figure 3.6 aligns the two DNA sequences in a matrix comparison. Apart from the coding region, there is virtually no conservation of DNA sequence between the two genes over the regions compared. The only exception is in the 5' flanking region (Figure 3.6B), where limited similarity is seen between regions rich in pyrimidine residues (on the upper strand), as well as three regions A, B and C (Figures 3.6B and 3.6C). The A regions share limited simlarity of 15/22 bp, while the B and C regions are more similar: The centers of these regions share 23/26 bp (B) and 26/31 bp (C). At the center of Region C is the sequence 5'-TGTCAAT-3', which is the VPE1 consensus sequence, identified in the promoter region of the vitellogenin genes from the two species (MacMorris *et al.*, 1994). The *C. elegans* regions A and B contain a 6/7 bp match to this same consensus, which has been shown to be important for high-level expression of the vitellogenin genes (Spieth *et al.*, 1985; MacMorris *et al.*, 1994). The relative 5' to 3' order of these three regions is conserved, although the distribution of the pyrimidine-rich segments is different. For example, in *C. briggsae*, regions A and B overlap and are not preceded by a pyrimidine-rich segment, while the opposite is true in *C. elegans*. Collectively, the regions are closer to the start of transcription in *C. elegans* by about 100 bp; it is curious that sequence C in *C. elegans* overlaps with the first exon, which suggests that it may not be a promoter element at all. Without evidence for alternative transcriptional initiation, a compelling argument for the maintenance of region C cannot be made. None of these sequences is present in the immediate upstream region of the C27H5.1 ORF. The functional significance of these regions, as well as the significance of the VPE1 elements, has not been experimentally tested.

3.3.5 Reporter gene expression is conserved between the two species

Evidence for functional conservation of *cis*-acting sequences comes from comparison of reporter gene expression. The *C. elegans unc-119* null mutants *ed3* and *ed4*, which contain nonsense mutations in Exon IV (Maduro and Pilgrim, 1995), were made transgenic for both a reporter gene fusion and a *C. elegans* genomic *unc-119* clone, such that animals carrying the transgenes have a wild type phenotype. Animals containing a *C. elegans* or *C. briggsae unc-119* gene fusion to βgal were prepared as described (Fire *et al.*, 1990). Photographs of differential interference contrast (DIC) microscopy of fixed and stained animals are shown in Figure 3.7.

Temporal expression is very similar, as transgene activity is seen very early in the embryo and continues through adulthood. Spatially, staining is restricted mainly to the nervous system. Staining of neuronal structures, such as the ventral nerve cord, the pre-anal and lumbar ganglia, and nerve ring, is readily seen in both fusion strains. There is striking similarity at higher magnifications as well. In panels (e) and (f), cells of the nerve ring and ventral nerve cord stain indistinguishably in placement and number of cells. Furthermore, some cells are seen to stain outside the nervous system, a phenomenon common to both constructs: Cells are visible anterior to the anterior bulb of the pharynx, where no neuronal cell bodies are found (White *et al.*, 1986). Both the *C. briggsae* and *C. elegans* photographs show staining similar to that previously reported for a *C. elegans unc-119* reporter gene (Maduro and Pilgrim, 1995).

To rule out differences due to amount of UNC-119 present in the fusions, both constructs were made to be very similar. The *C. briggsae* fusion contains ~620 bp of upstream DNA and 130 aa of UNC-119; the *C. elegans* fusion contains ~1160 bp of promoter and 101 aa of UNC-119. Both contain the conserved regions A, B and C. The similar expression pattern is consistent with the maintenance of the specificity of these *cis*-acting regions, and hence transcriptional control. Alternatively, transcription may occur at a basal level throughout the animal, and visualization of specific nervous system staining would arise from post-translational regulation, since the amino-terminal half of UNC-119 is present in both fusions. Such regulation might be maintained due to the high degree of amino acid similarity; however, when the same *C. elegans unc-119* promoter (without any coding sequences) is used to drive expression of another reporter, the green fluorescent protein (GFP) from *Aequoria victoria* (Chalfie *et al.*, 1994), specific nervous system expression still occurs (data not shown), suggesting regulation of *unc-119::*reporter expression is primarily transcriptional.

It is possible that similar neuronal expression at the transcriptional level is due to smaller promoter elements scattered throughout the control region, but the conservation of position and sequence of the A, B and C elements, as well as the existence of a known transcriptional regulatory element at the core of regions B and C, make these good candidates for putative neuronal control elements.

3.3.6 Conclusion

The *unc-119* genes from *C. elegans* and *C. briggsae* provide another example in which conservation of coding sequence, coupled with maintenance of only a small number of control elements, allow similarity of function and expression between two species that are separated by tens of millions of years of evolution. Here we have demonstrated that inter-species sequence comparisons provide a means by which homologous genes and candidate promoter elements may be identified.

Acknowledgements

We wish to thank Terry Snutch and David Baillie for providing the *C. briggsae* genomic library, and Andrew Fire for donating fusion vectors. This work was supported by a grant from the Natural Science and Engineering Research Council.

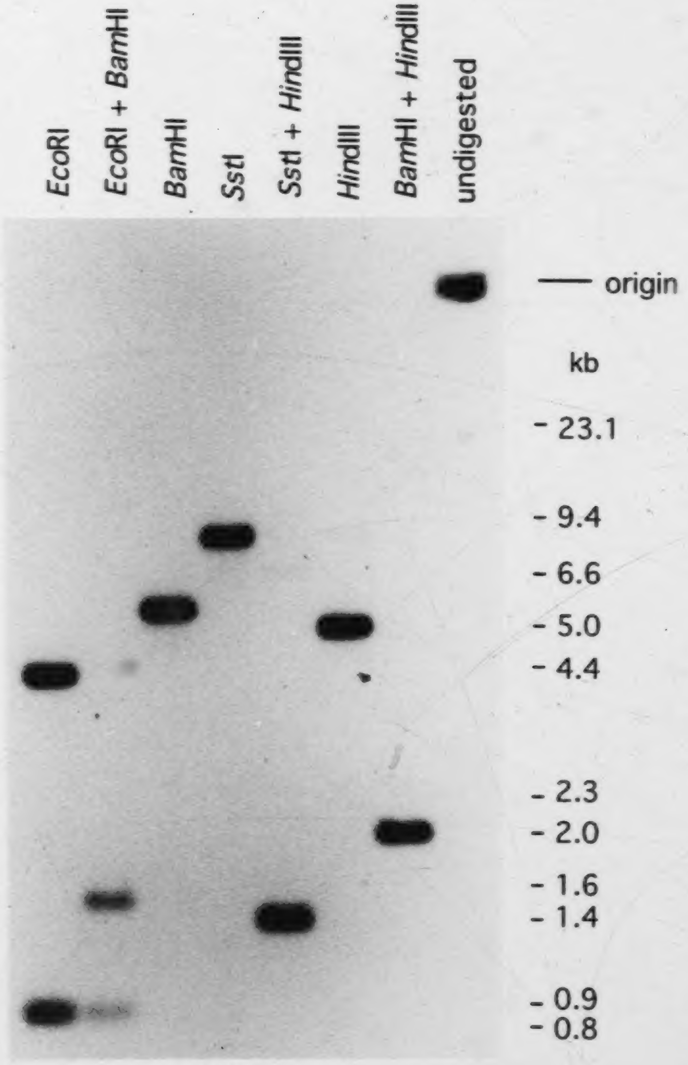


Figure 3.1 Southern (1975) analysis of restriction-digested pDP#MMCb1.

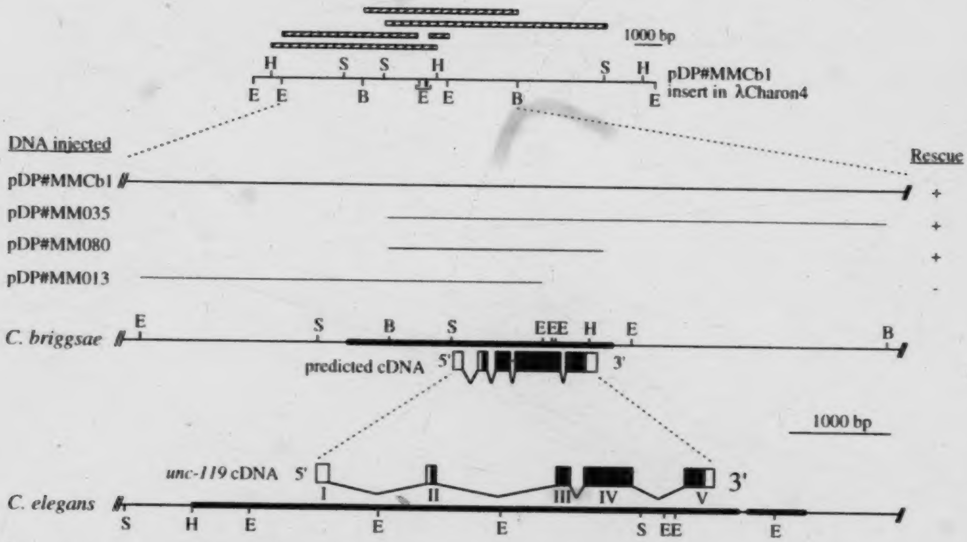


Figure 3.2 Restriction map of the *C. briggsae* insertion from the λCharon4 clone pDP#MMCb1, rescuing ability of subclones, and the predicted cDNA aligned with the *C. elegans* counterpart. Fragments that cross-hybridize with the *unc-119* cDNA (from Figure 3.1) are shown above the map as shaded boxes. Transgenic animals carrying the various clones were constructed and phenotypically analyzed as previously described (Maduro and Pilgrim, 1995). Rescue is denoted as "+", while failure to rescue is shown as "-". The clone pDP#MM080 is a fusion of *C. briggsae unc-119* at an *Nde*I site in the 3' end of the coding region to a *lacZ* reporter gene containing the *unc-54* 3' untranslated region (UTR) (Fire, 1990) which excludes only the last amino acid in the ORF. Rescue by pDP#MM035 was only ascertained on the basis of rescue of the locomotory defect. The thicker regions on the bottom maps respresent those areas that have been sequenced. The *C. elegans unc-119* sequence, which has been published (Maduro and Pilgrim, 1995), is available through GenBank (Accession No. U32854). Abbreviations: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; S, *Ssf*I.

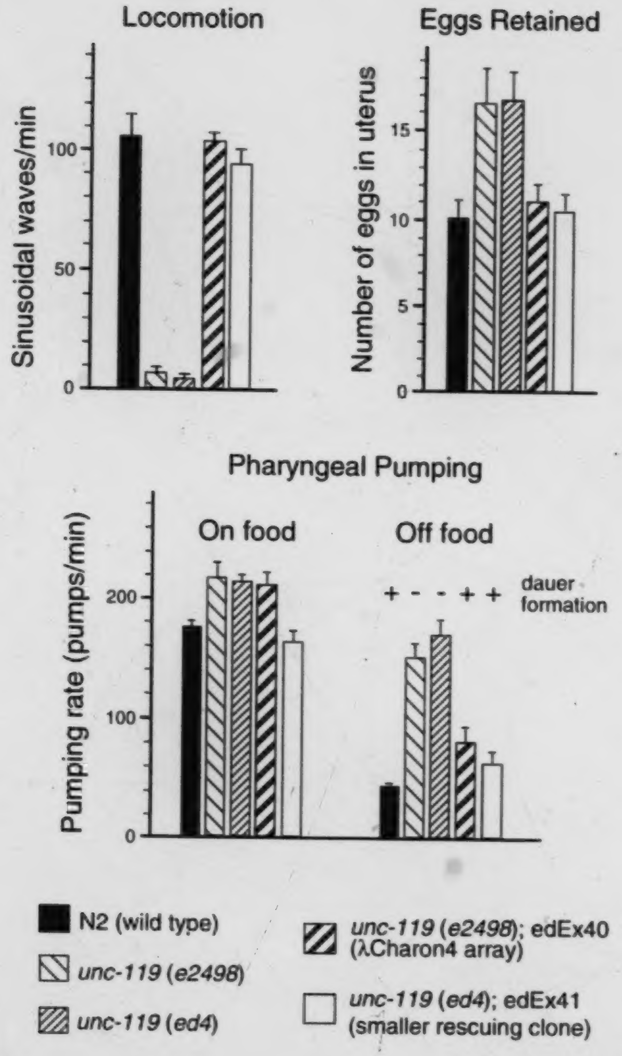


Figure 3.3 Quantitation of defects in wild type, *unc-119* mutants and mutants carrying the putative *C. briggsae unc-119* gene on transgenic arrays. The *e2498* mutation results from a transposon insertion 3' to Exon IV, and *ed4* results from a nonsense mutation in Exon IV (Maduro and Pilgrim, 1995). Bars show the standard error of the mean (SEM). Data collection and the dauer-formation assay were performed as described (Maduro and Pilgrim, 1995). The array *edEx40* contains pDP#MMCb1 and pRF4; *edEx41* contains pDP#MM080 alone.

```
             5'-gtctcatagattacatgagtaattggtctctatttctgaaagttccgatctttcaaatgttcagttaacagcagtgttcacaca -1001
tatttgaccagccttcggattcgaaataaacaaactgtcttttccaaattgtttccttccgttttcaagtttgaatgactcatgtaagggtgcatccaaa  -901
aagaagaattttcgaactagcgatgctgaagtgtggctagagcaaactagacagtcctaacagagccatttcgccagtttcgtaaagttaccatgaaac  -801
ctgcaatagaagggtcccaactctgttaagtttgctagtttttcttatctcctgtgcaagatatcaaaaatgtgtcaactgataaaatttaatgatgtca  -701
tgatttatcagttgacaacttttatatctttcacaggaacagagataatatggctcaaagtgcaaacgtcctatggatcctatgcttgcgctttgagcc  -601
aatttatccaagtccttgtaaaaagtattcgaaaattgttaacggataaaatgtttattataatatcaaaaacaatttgtcagttgaccacttttgata  -501
tatttgacagaaaccgggatgaattggctcaaagttagggcgcccttctattacagggtttctgataacaaaccggttattaactcccaacaagggatga  -401
tttcaattcatcatgctcaaattgacccaaattaagttacatgacaaattcatcgcccttttcaactcttcttggctcatcatctgttattctgttccta  -301
ttctgtctgcaccctataccccttgcatactctctcgtcattcctttcttggatagtgtcttcttctccccagctctgctacttctatgacttgcgccc  -201
gctgcttttccgctcgttcctctcctctctgacgtcttcttctgcttcttttctccttcctttttcccttctcaatcttcctctttttcccatttctgtc  -101
aatcattcgaagaagaagaagaagaagagccctcattcatttcatttttttttctgtcggtgtgtgctgctcggttaagagtgagctctctattccacgt    -1
CTTCTTCTTTTTCTTCTTCGATTCGAATCAATCACTCCACCAAAACGCATTCGTTTTTGGGATTCACCCCGCGGTTCGCAACAGgtttctttttcaaatt   100
attagcgttataaatagaaaaatgtggagtttcaataaaaatgataatttacaaagtgtatttgattacatgtactcaaaaaggttgaaattttcata   200
ccagttttccggaaatccatctgatatcattatcgtatttttcttttaaaaatgttttcaaaaaaaaacaaaatatagctggttatttggcaccctcta   300

                                                                               M  K  A  E  Q  Q  Q
attaccattttcctgtcacaccacactcttttccttcttccacttctttccgttttcagCCGCTTCCAACCAAACCGATATGAAAGCCGAGCAACAACAA   400
S  I  P  P  G  S  A  T  F  P  S  Q
TCGATTCCACCCGGCTCGGCGACCTTCCCGTCGCAGgtgagactcagaaactagagaaccgctcaactactcttgatatcgatttcatctttctctttct   500
                   M  P  R  P  P  P  S  T  E  Q  G  I  T  T  E  S  E  L  A  K  K  A  Q  I  T  P  N
ttgtgacttcatattccagATGCCACGGCCACCACCAAGCACCGAACAAGGAATCACAACGGAATCGGAGCTTGCGAAGAAAGCTCAAATCACTCCGAAC   600
D  V  L  A  L  P  G  I  T  Q  G                                              F  L  C  S  P  S  A  N
GACGTTTTAGCACTTCCGGGAATCACTCAAGgtatcttgctccctttggatttcagtttaaaacataaatttagGATTCTTATGCTCCCCATCTGCGAAC   700
I  Y  N  I  E  F  T  K  F  Q  I  R  D  L  D  T  E  Q  V  L  F  E  I  A  K  P  E  N  D  Q  E  N  D  E
ATCTATAACATCGAGTTCACCAAGTTCCAAATCCGTGATCTGGACACTGAGCAAGTGCTTTTCGAGATCGCCAAACCGGAGAACGATCAGGAGAATGATG   800
  S  P  Q  E  S  A  R  Y  V  R  Y  R  F  A  P  N  F  L  K  L  K  T  V  G  A  T  V  E  F  K  V  G  D
AGTCGCCACAGGAGTCGGCAAGATACGTGCGTTATAGATTTGCTCCAAACTTTTTGAAACTCAAAACGGTCGGAGCAACTGTGGAATTCAAAGTAGGAGA   900
 I  P  I  H  H  F  R  M  I  E  R  H  F  F  F  D  R  L  L  K  C  F  D  F  E  F  G  F  C  I  P  N  S
CATCCCAATCCATCATTTCCGAATGATCGAACGTCACTTCTTCTTTGATCGCCTTCTGAAGTGTTTTGACTTTGAATTCGGATTCTGTATTCCGAATTCA  1000
```

Figure 3.4 The genomic DNA sequence of *C. briggsae unc-119*. Predicted exons are shown in upper case, with conceptual translation above the sequence in single letter code above the first base of each codon. The start and stop codons and the putative polyadenylation sequence are underlined.

```
R  N  N  C  E  H  I  Y  E  F  P  Q  L  S  Q  Q  L  M
CGAAACAACTGTGAACATATCTATGAGTTCCCTCAACTCTCTCAACAACTCAgtgagtcattattctaaaagtacaattcaaaagactaatctcttttt  1100
     D  D  M  I  N  N  P  N  E  T  R  S  D  S  F  Y  F  V  D  N  K  L  V  M  H  N  K  A  D  Y  S  Y
cagTGGACGACATGATCAACAATCCAAACGAGACTCGTTCTGACAGCTTCTATTTCGTCGATAACAAACTCGTCATGCACAACAAAGCCGACTACTCATA  1200
 D  A
TGATGCATAAATATTTAATACAAAATGTTCTGGATAATTATTCTGTCGAATAGAAAAAACTCCAATGTGATTAAATTCCAATATTCCTGTCTGTTTTGTC  1300
CTTCCCTTCCCTCTCTGTCATGCATTCTAAGCTTTTCAGTTCCCCCCTTGTTTCTATATTTTTTCGCTGTCCTGTCACACTCGCTAAAAACACTAATCAC  1400
ACGGAAATCTGTTTTCAATAAAAACTCCAACTTTAactcattttcaatttcaactgaaagatttttcattagagaatgtccagcgaggaggatacattg  1500
tacccatcaaacttgcgcaaagatttattgtaggttttcagtactttttgagggagaaaaagtaatctcactaaaactttaaaaaaactgatactggcac  1600
tcaaggatattgttctttttgagggcttccgggaaaatttgaaatttaaaggaaggaaaaaatcattgaaaagatttcaac-3'  1681
```

Figure 3.4, continued

```
C. elegans   MKAEQQQQSIAPGSATFPSQMPRPPPVTEQAITTEAELLAKNQITPNDVL
C. briggsae  ········−·················S···G····S··AK·A········
C27H5.1      MATT  (C27H5.1 is shorter than the other two proteins)

                                 ‡         ‡     ‡
C. elegans   ALPGITQGFLCSPSANVYNIEFTKFQIRDLDTGHVLFEIAKPENETEENL
C. briggsae  ················I···············EQ···········Q−··D
C27H5.1      ·TRHQDSKLSEKAESILAGFKLNWMNL··AE··K··WQST−−−−−−−−−−

             ◊   ‡     ‡          ‡                        ‡
C. elegans   QAQAESARYVRYRFAPNFLKLKTVGATVEFKVGDVPITHFRMIERHFFKD
C. briggsae  ESPQ······························I··H··········F·
C27H5.1      EDM·DPK·EHKAHVPK·L··CR··SREIN·−TSS·K·EK··LEQ·VYL·G

                ‡ ‡  ‡    ◊                                ‡
C. elegans   RLLKCFDFEFGFCMPNSRNNCEHIYEFPQLSQQLMDDMINNPNETRSDSF
C. briggsae  ·············I····································
C27H5.1      TIIEEWY·D···VI·D·T·TWQNMI·AAPE··MFPPSVLSGNVVVET−L·

                ◊       ‡
C. elegans   YFVENKLVMHNKADYSYDA*
C. briggsae  ···D················
C27H5.1      ·DGDL−··STSRVRLY··*
```

Figure 3.5 Alignment of the *C. elegans* and *C. briggsae* predicted UNC-119 protein sequences, and the predicted product of the open reading frame C27H5.1 from *C. elegans*. •; amino acid identical to *C. elegans*; -, gap introduced to optimize alignment; ◊, amino acid identical between C27H5.1 and *C. briggsae*; ‡, conservation of charged or hydrophobic amino acid; *, stop codon. The GenBank Accession number for C27H5.1 is U14635.

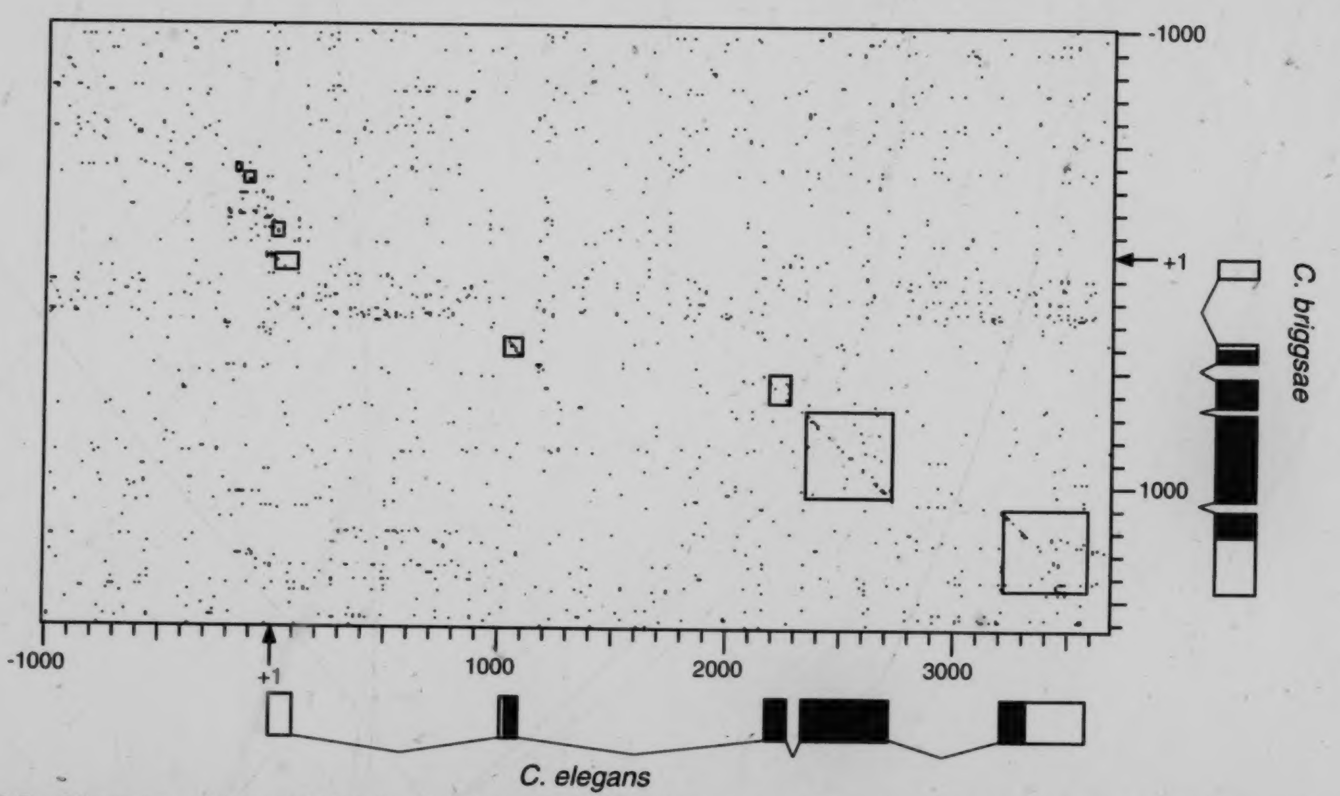


Figure 3.6 Alignment of *unc-119* DNA sequences from *C. elegans* and *C. briggsae.* Numbers refer to base pairs of DNA sequences as shown in Figure 4 (*C. briggsae*) and Maduro and Pilgrim (1995) (*C. elegans*). ***A***, Dot matrix alignment of entire coding and flanking regions. Boxed regions along the axis refer to exons of the genes; shaded boxes correspond to the protein coding regions of the transcript. Boxed regions within the matrix correspond to the overlap of exon sequences from the two species, as well as conserved 5' flanking sequences. Alignment was produced using DNA Strider 1.2 (Marck, 1988), using Stringency of 7, Window of 7 and Scale of 10.

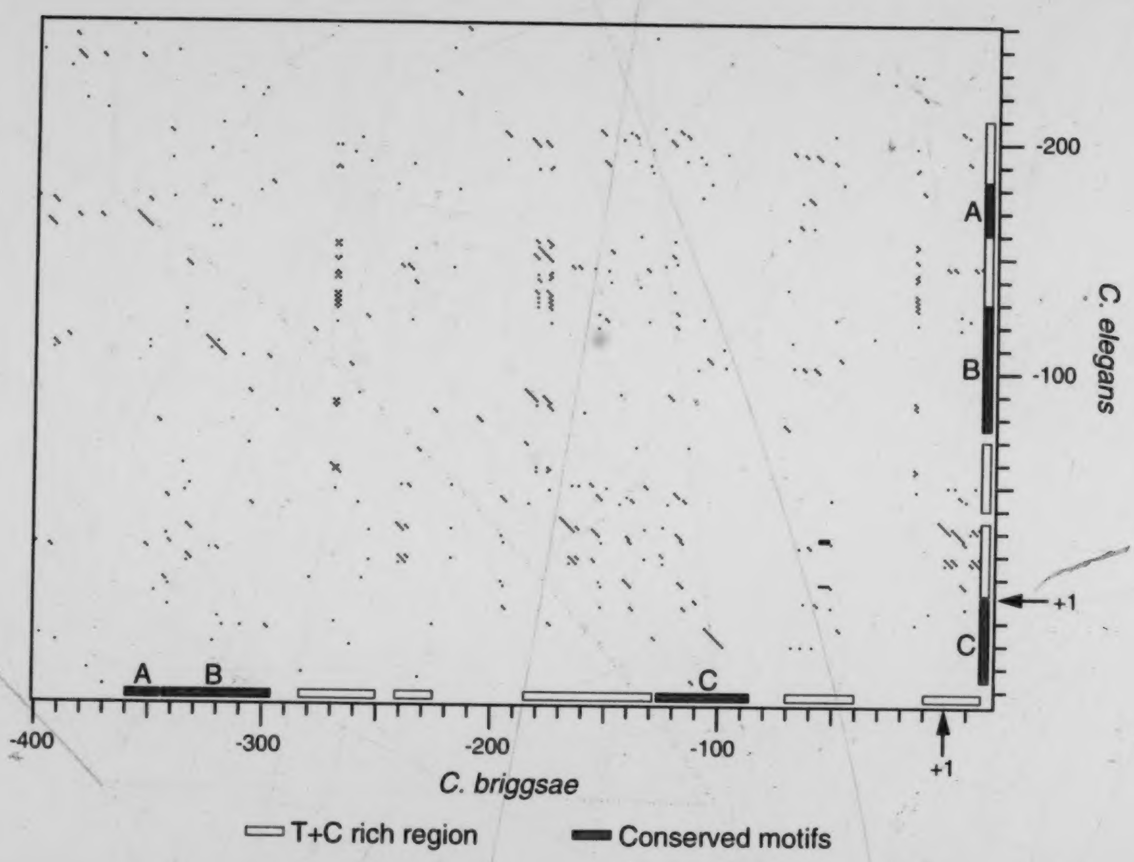


Figure 3.6, continued. ***B***, Enlargement of 5' flanking regions. Conserved regions A, B and C (see below) are highlighted along the axes by shaded boxes, pyrimidine-rich regions are highlighted by open boxes. Alignment was produced using DNA Strider 1.2 (Marck, 1988), using Stringency of 5, Window of 5 and Scale of 1.

Sequence A

```
                 ** * * ****** *    ****
C. elegans  -181 TGTCTATTTCATCAC-AAATTCA
C. briggsae -359 TGACAAATTCATCGCCCTTTTCA
```

Sequence B

```
                  * **   *  **** ************* *** *** * * * ** ***
C. elegans  -128 CTCTCTCTTTCTCTTTGCTCATCATCTGTCATTTTGTCCGT-TCCTCTCT
C. briggsae -341 TTTTCAACTCTTCTTGGCTCATCATCTGTTATTCTGTTCCTATTCTGTCT
```

Sequence C

```
                  ***  * **** **** ************* * ** ** *
C. elegans    -4 TTTCGCCATTTT-CCATCTCTGTCAATCATTACGGACGACG
C. briggsae -124 CTTCCTCTTTTTCCCATTTCTGTCAATCATT-C-GAAGAAG
```

Sequences B + C

```
C. elegans B   CTTTCTCTTTGCTC-ATCATCTGTCATT--TTGTCCGTTC
C. briggsae B  ACTCTTCTTGGCTC-ATCATCTGTTATTC-TGTTCCTATT
C. elegans C   TTTCGCCATTTTCC-ATC-TCTGTCAATCATTA-CGGACG
C. briggsae C  CTTCCTCTTTTTCCCATT-TCTGTCAATCATT---CGAAG

CONSENSUS      YTTCYTCTTTNYYC-ATCATCTGTCANTCATTNTCCGAYN
```

Figure 3.6, continued. ***C***, Similarity between the 5' flanking regions of the *unc-119* gene. Numbers give distances 5' to the start of transcription (+1). Asterisks indicate conserved base pairs, dashes where gaps were introduced to maximize alignment. Underlined sequences highlight 6 of 7 base pair match to VPE1 *C. elegans* promoter element (MacMorris *et al.*, 1994). Bold letters in region B match the 'E-box' sequence 5'-CATCTG-3' (Krause et al, 1994). An alignment of sequences B and C, along with a consensus, is shown immediately below.

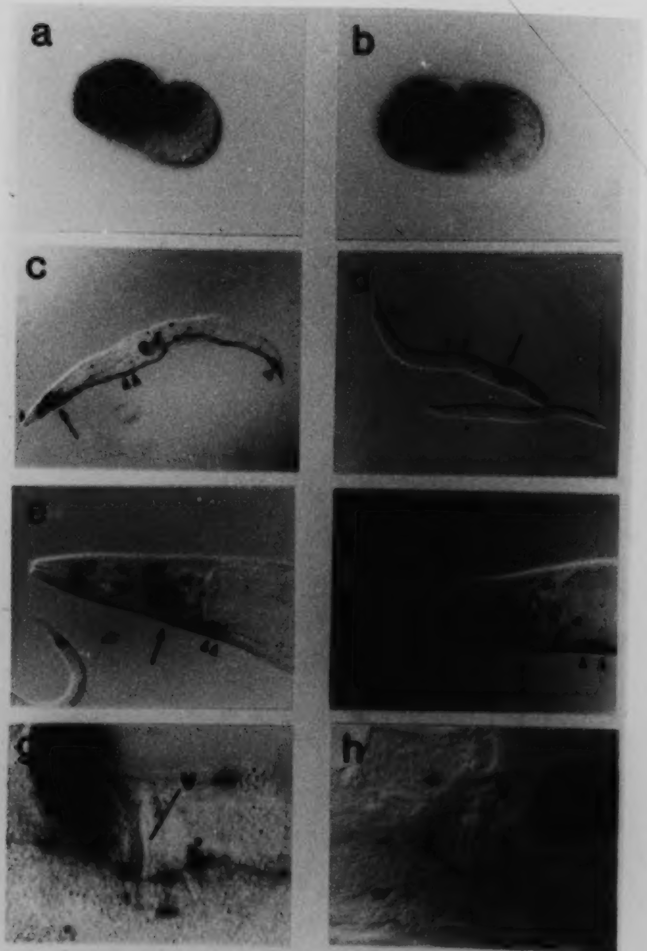


Figure 3.7 X-gal staining of *C. elegans* hermaphrodites transgenic for *unc-119::*βgal fusions. Left column, *unc-119* (*ed4*); *edEx42* (*C. briggsae* fusion; *edEx42* contains the fusion plasmid pDP#MM082 and the rescuing clone pDP#MM016); right column [except (**h**)], *unc-119*(*ed3*); *edEx43* (*C. elegans* fusion; *edEx43* contains the fusion plasmid pDP#MM081 and pDP#MM016); (**h**), wild type animal transgenic for pRF4 (rol-6D marker) and pDP#MM082. Despite the nuclear localizing signal (NLS) present in the reporter gene constructs, staining is still visible outside of nuclei (as verified by DAPI staining; data not shown). Panels (**a**) and (**b**), comma-stage embryos showing anteriorly-localized staining in presumptive neuronal cell precursors; (**c**) adult and (**d**) fourth larval stage animals, showing staining in the nerve ring (arrows), pre-anal ganglia (open arrowheads) and ventral nerve cord (filled arrowheads); (**e**) and (**f**), adults showing staining in the nerve ring (arrows) and ventral nerve cord (filled arrowheads); (**g**) and (**h**), ventral view of hermaphrodite vulva (V) showing the ventral nerve cord and a VC neuron (*). A diffusely staining early embryo (E) is visible in each panel.

3.4 Bibliography

Berks, M., and the *C. elegans* genome mapping and sequencing consortium. (1995). The *C. elegans* genome sequencing project. Genome Res. *5*, 99-104.

Burns, N., Grimwade, B., Ross-Macdonald, P. B., Choi, E.-Y., Finberg, K., Roeder, G. S., and Snyder, M. (1994). Large-scale analysis of gene expression, protein localization, and gene disruption in *Saccharomyces cerevisiae.* Genes Dev. *8*, 1087-1105.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. Science *263*, 802-805.

Fire, A., Harrison, S. W., and Dixon, D. (1990). A modular set of *lacZ* fusion vectors for studying gene expression in *Caenorhabditis elegans.* Gene *93*, 189-198.

Frohman, M. A., Dush, M. K., and Martin, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specifc oligonucleotide primer. PNAS (USA) *85*, 8998-9002.

Goodwin, E. B., Okkema, P. G., Evans, T. C., and Kimble, J. (1993). Translational regulation of *tra-2* by its 3' untranslated region controls sexual identity in *C. elegans.* Cell *75*, 329-339.

Heschl, M. F. P., and Baillie, D. L. (1990). Functional elements and domains inferred from sequence comparisons of a heat shock gene in two nematodes. J. Mol. Evol. *31*, 3-9.

Hodgkin J., Plasterk, R. H. A., and Waterston, R. H. (1995). The nematode *Caenorhabditis elegans* and its genome. Science *270*, 410-414.

Jones, D., and Candido, E. P. M. (1993). Novel ubiquitin-like ribosomal protein fusion genes from the nematodes *Caenorhabditis elegans* and *Caenorhabditis briggsae*. J. Biol. Chem. *268*, 19545-19551.

Kennedy, B. P., Aamodt, E. J., Allen, F. L., Chung, M. A., Heschl, M. F. P., and McGhee, J. D. (1993). The gut esterase gene (*ges-1*) from the nematodes *Caenorhabditis elegans* and *Caenorhabditis briggsae*. J. Mol. Biol. *229*, 890-908.

Krause, M. (1995). Transcription and translation. Methods Cell Biol. *48*, 483-512.

Krause, M., Harrison, S. W., Xu, S.-Q., Chen, L., and Fire, A. (1994). Elements regulating cell- and stage-specific expression of the *C. elegans* MyoD family homolog *hlh-1*. Dev. Biol. *166*, 133-148.

Kuwabara, P. E., and Shah, S. (1994). Cloning by synteny: identifying *C. briggsae* homologues of *C. elegans* genes. Nucleic Acids Res. *22*, 4414-4418.

Lee, Y. H., Huang, X. Y., Hirsh, D., Fox, G. E., and Hecht, R. M. (1992). Conservation of gene organization and trans-splicing in the glyceraldehyde-3-phosphate dehydrogenase-encoding genes of *Caenorhabditis briggsae*. Gene *121*, 227-235.

MacMorris, M., Spieth, J., Madej, C., Lea, K., and Blumenthal, T. (1994). Analysis of the VPE sequences in the *Caenorhabditis elegans vit-2* promoter with extrachromosomal tandem array-containing strains. Mol. Cell. Biol. *14*, 484-491.

Maduro, M., and Pilgrim, D. (1995). Identification and cloning of *unc-119*, a gene expressed in the *Caenorhabditis elegans* nervous system. Genetics *141*, 977-988.

Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. EMBO J. *10*, 3959-3970.

Marck, C. (1988). "DNA Strider": a "C" program for the fast analysis of DNA and Protein sequences on the Apple Macintosh family of computers. Nucleic Acids Res. *16*, 1829-1836.

Pilgrim, D., McGregor, A., Jäckle, P., Johnson, T., and Hansen, D. (1995). The *C. elegans* sex determining gene *fem-2* encodes a putative protein phosphatase. Mol. Biol. Cell *6*, 1159-1171.

Prasad, S. S., and Baillie, D. L. (1989). Evolutionarily conserved coding sequences in the *dpy-20-unc-22* region of *Caenorhabditis elegans*. Genomics *5*, 185-198.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, NY.

Snutch, T. P. (1984). Ph. D. Thesis, Simon Fraser University, Burnaby, BC.

Southern, E. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. *98*, 503-517.

Spieth, J., Denison, K., Kirtland, S., Cane, J., and Blumenthal, T. (1985). The *C. elegans* vitellogenin genes: short sequence repeats in the promoter regions and homology to the vertebrate genes. Nucleic Acids Res. *13*, 5283-5295.

Stroeher, V., Kennedy, B. P., Millen, K. J., Schroeder, D. F., Hawkins, M. G., Goszczynski, B., and McGhee, J. D. (1994). DNA-protein interactions in the *Caenorhabditis elegans* embryo: oocyte and embryonic factors that bind to the promoter of the gut-specific *ges-1* gene. Dev. Biol. *163*, 367-380.

White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1986). The structure of the nervous system of *Caenorhabditis elegans*. Phil. Trans. R. Soc. Lond. B Biol. Sci. *314*, 1-340.

Wood, W. B., ed. (1988). The Nematode *Caenorhabditis elegans*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Zucker-Aprison, E., and Blumenthal, T. (1989). Potential regulatory element of nematode vitellogenin genes revealed by interspecies sequence comparison. J. Mol Evol. *28*, 487-496.

3.5 Addendum

Subsequent to the publication of this chapter, additional experiments were performed. The first of these is an analysis of *unc-119* messenger RNA in mixed-stage preparations of *C. briggsae* and *C. elegans*, that confirms the existence of appropriately sized transcripts. The second is a description of the expression pattern of a green fluorescent protein (GFP) reporter fusion of the *C. briggsae* gene expressed in *C. elegans*. Lastly, a brief description of a homologue to C27H5.1 is given.

3.6 Methods

3.6.1 Preparation of total RNA

Total RNA was prepared from crowded 6-cm plates of wild type *C. elegans* and *C. briggsae* as follows. Worms were washed off plates with 0.1M NaCl, and rinsed to remove excess bacteria. To 150 µL of worm suspension, 400 µL of GuEST buffer, 400 µL of PCI (phenol:chloroform:isoamyl alcohol 25:24:1) and 2 µL of β-mercaptoethanol were added. The mixture was ground for one minute using a tissue homogenizer and then transferred to an eppendorf tube. The homogenizer was rinsed with 200 µL of GuEST and 200 µL PCI, which was added to the first mixture. The tube was spun at full speed in a benchtop microfuge for 3 minutes. Approximately 600 µL of aqueous phase were retrieved to a new tube, 60 µL of 3M NaOAc were added, and this was extracted three times with PCI and once with chloroform. Two volumes of 100% ethanol were added to precipitate the RNA, which was pelleted by centrifugation, rinsed in 80% ethanol, and dissolved in 30 µL of DEPC-treated H_20. Approximately 10 µg of total RNA were obtained per 6-cm plate.

3.6.2 *Northern hybridization*

Formamide mini-gels prepared by standard protocols (Sambrook *et al.*, 1989) were run of 10 µg total RNA at 80 V for two hours. RNA was blotted onto Hybond N nylon membrane (Amersham) by alkaline capillary transfer as described in Sambrook *et al.* (1989). An RNA molecular size standard ladder was used (Gibco-BRL).

To detect the mRNAs, run-off anti-sense transcripts containing α-^{32}P-rUTP were prepared from linearized pBluescript KS- clones using the T3 RNA polymerase supplied with an *in vitro* transcription kit (Promega). For the *C. elegans* gene, a cDNA fragment clone containing only the UNC-119 coding sequences was used (cDNA7 linearized with *Xba*I); for the *C. briggsae* gene, a 1.3 kb *Hin*DIII-*Sst*I genomic fragment (pDP#MM078) was linearized with *Eco*RV; the run-off transcript is predicted to be complementary to exons III, IV and V of the *C. briggsae* gene (see Figure 3.8). Hybridization was performed in 50-mL Falcon tubes at 60°C in a 10 mL solution made up of 5 mL formamide, 2.5 mL 20×SSPE, 0.5 mL 100×Denhardt's reagent, 0.5 g SDS, and DEPC-treated water. After overnight hybridization, membranes were rinsed 3× at room temperature for 30 minutes in 2×SSPE/1%SDS, then for 10 minutes at 65°C in 0.2×SSPE/1% SDS. Blots were exposed to film for four days at -60°C with two intensifying screens.

3.6.3 *C. briggsae unc-119::GFP fusion plasmid*

The *C. briggsae* GFP fusion pUGF15 was made by removing a *Hin*DIII-*Pst*I fragment from pDP#MM070 (see §3.2.3) and cloning into the same sites in pPD95.77 (A. Fire, J. Ahnn, G. Seydoux and S. Xu, pers. comm.). Plasmid pUGF15 is predicted to express 130 aa of CbUNC-119 fused to the S65C 'enhanced' GFP.

3.7 Northern Blot Analysis of *unc-119*

The sequences of the genomic *unc-119* gene from both species, and the 5'RACE products, predict messages of 986 bases (*C. elegans*) and 983 bases (*C. briggsae*), without consideration of polyadenylation. A preliminary attempt to demonstrate the presence of these transcripts in total RNA was performed on both species.

Northern blots from *C. elegans* and *C. briggsae* are shown in Figure 3.9. In both, a weak cross-hybridizing band of approximately 1000 bases is seen. In the *C. briggsae* lane, additional cross-hybridization is seen at positions consistent with the migration of the ribosomal RNA (shown with asterisks in Figure 3.9), suggesting these bands might be nonspecific.

These data corroborate the existence of a single *unc-119* message in *C. elegans*, consistent with the results of 5'RACE and the sequences of the cDNA clones; the resolution of the gels, however, may obscure an alternatively spliced mRNA of similar size. While such an alternate transcript may be produced, there is good evidence that only a single transcript is required: The mini-gene fusion pDP#MM051, which fuses the promoter with the full-length cDNA, can completely rescue the *unc-119* mutant defects (Chapter 2).

3.8 Reporter gene fusion of *C. briggsae unc-119* with GFP

The availability of vectors encoding the green fluorescent protein (GFP) from the jellyfish *Aequoria victoria* (Chalfie *et al.*, 1994) provided an opportunity to examine the transgene expression pattern of *C. briggsae unc-119* in *C. elegans*. This allowed observation of expression in living animals, and at higher resolution than is possible with the *lacZ* reporter used previously (see Chapter 4).

A GFP reporter was made with the same upstream sequences as those used to make the *C. briggsae unc-119*::lacZ fusion described earlier. The fusion plasmid, pUGF15, is predicted to encode 130 aa of *C. briggsae* UNC-119 (see §3.6.3). *C. elegans* animals homozygous for *unc-119*(*ed3*) were made transgenic for pUGF15 and pDP#MM016B, a plasmid containing the wild type *unc-119* gene, by germline transformation (Mello *et al.*, 1991).

Transgenic animals were examined by fluorescence microscopy. The reporter had the same nervous system expression pattern as the *lacZ* fusion, with the following exceptions. First, fluorescence was found to occur throughout axons, rather than in nuclei. This is attributed to the absence of the nuclear localization signal (NLS) that was found in the original *lacZ* expression vectors (Fire *et al.*, 1990).

Second, unusual features of the expression became apparent with the use of GFP as a reporter. The localization of the fluorescence appeared punctate, rather than diffuse, within axons (Figure 3.10A). Large clusters of GFP were visible in the ventral nerve cord and nerve ring, suggesting these might be cell bodies. However, labeling of nuclei with DAPI showed that these clusters were cytoplasmic (Figures 3.10D-F).

In addition to the punctate nervous system fluorescence, signal appeared in punctate regular arrays in the head (Figure 3.10C). These occurred in the same plane of focus as muscle, and exhibited the same regular pattern as fluorescently labeled muscle dense bodies (Figures 3.10G,H).

A closer examination of the staining pattern of the original *C. briggsae lacZ* fusion revealed similar artifacts that were previously unnoticed; presumably, the nuclear localization had reduced the amount of cytoplamic fusion protein that could exhibit punctate staining.

The implications of these localization patterns will be discussed in the addendum to Chapter 4.

3.9 Similarity of C27H5.1 to a vertebrate protein

The amino acid sequence of UNC-119 was found to be weakly similar to C27H5.1, a predicted gene from the *C. elegans* sequencing project (§2.3.4 and Figure 3.5). A database search using this ORF revealed similarity to the ∂ subunit of rod and cone phosphodiesterase (D. Pilgrim, pers. comm.; Florio *et al.*, 1996). The function of this subunit is not known, although its association with retinal specific phosphodiesterases suggests a role in phototransduction (Florio *et al.*, 1996). The implications of this similarity will be discussed in Chapter 6.

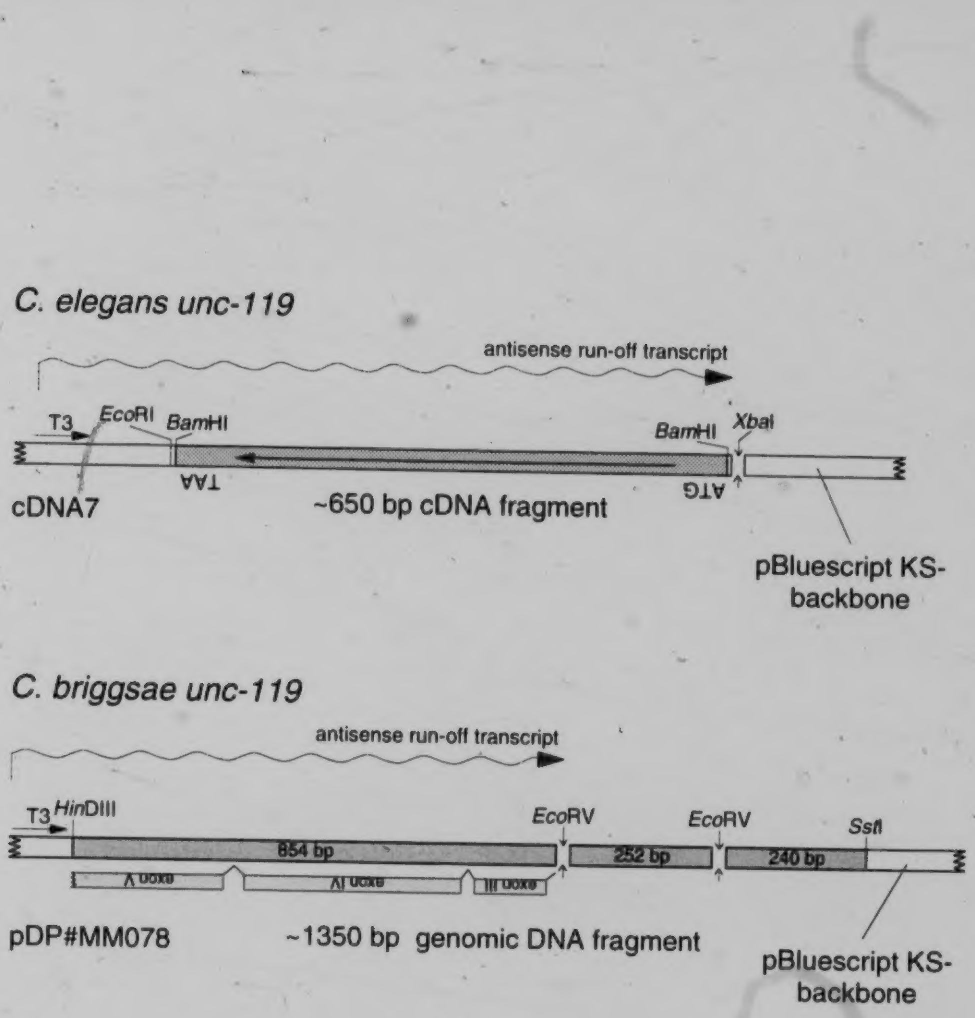
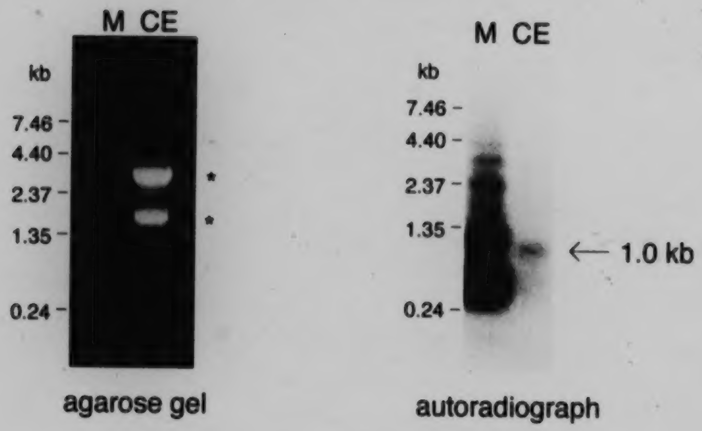


Figure 3.8 Run-off templates for synthesis of *unc-119* specific antisense RNA. Top, *C. elegans Bam*HI fragment of cDNA, which contains the entire *unc-119* coding region (clone cDNA7). Bottom, *C. briggsae Hin*DIII-*Sst*I genomic clone pDP#MM078, which allows run-off synthesis of an RNA that is complementary to *unc-119* exons III, IV and most of V. Both antisense RNAs were synthesized *in vitro* using T3 RNA polymerase. The shaded regions show the actual insertion in pBluescript KS-; gaps show the sites digested by the enzyme used to linearize the template.

C. elegans Northern Blot



C. briggsae Northern Blot

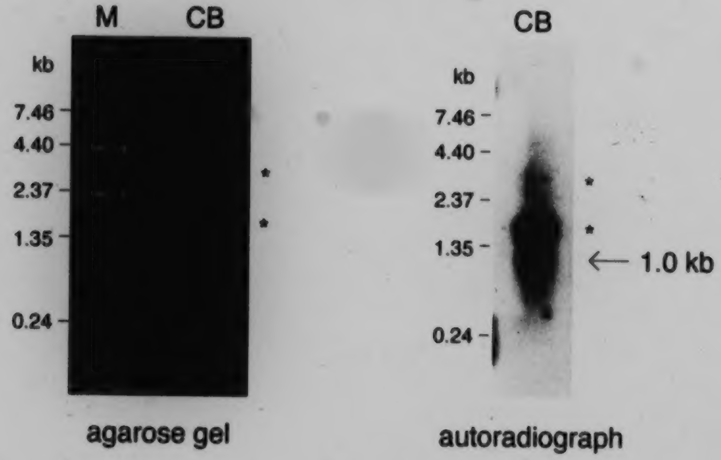


Figure 3.9 Northern hybridization analysis of *unc-119* in mixed-stage total RNA preparations from *C. elegans* and *C. briggsae*. The *unc-119* transcript bands are shown with an arrow. The marker lane, which shows cross-hybridization in the *C. elegans* blot, was removed from the *C. briggsae* membrane prior to probing. The *C. elegans* blot was exposed to film for ten days; the *C. briggsae* blot was exposed to film for five days. M, marker lane; CE, total *C. elegans* RNA; CB, total *C. briggsae* RNA; *, rRNA band.

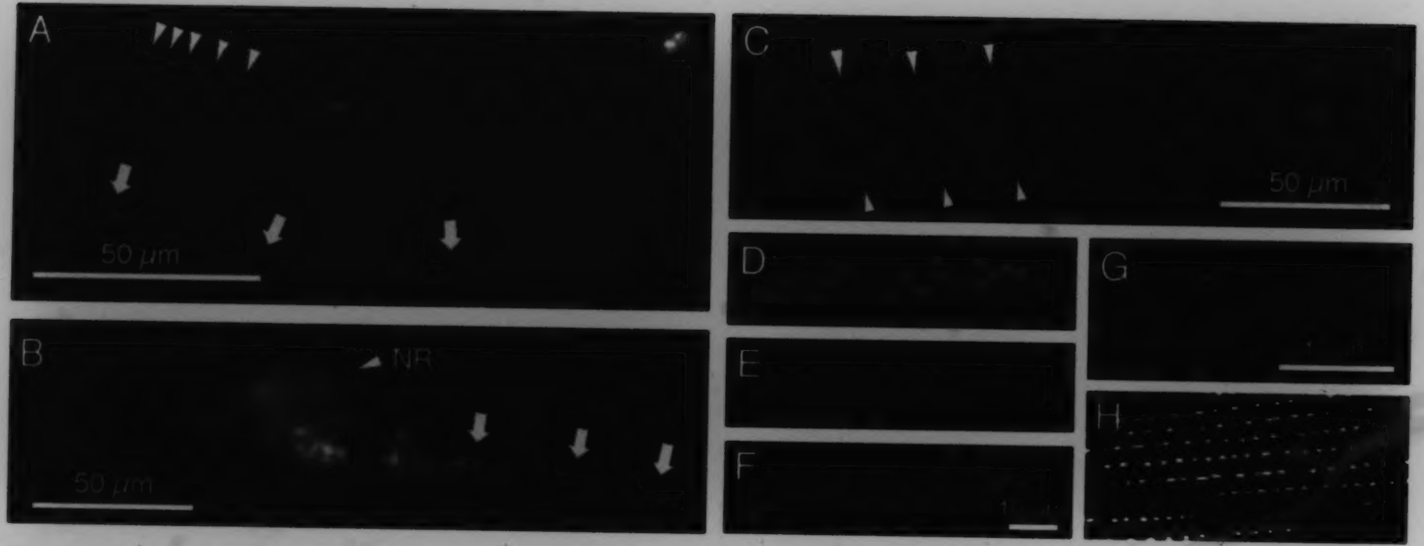


Figure 3.10 *unc-119(ed3)* animals transgenic for an array carrying pUGF15 (*C. briggsae unc-119*::GFP fusion) and pDP#MM016B (an *unc-119* rescuing clone). **A**, Punctate GFP fluorescence (arrowheads) is seen along processes. The ventral nerve cord, just out of the plane of focus, also expresses GFP (arrows). **B**, In the vicinity of the nerve ring (NR), bright punctate localization can be seen, which continues to the ventral nerve cord (arrows). **C**, In the head, longitudinal zones of regularly spaced structures fluoresce. **D**, **E**, and **F**, ventral nerve cord punctate staining is not nuclear. DAPI fluorescence (blue) indicates nuclei (images D and F), and green is the GFP signal (images E and F). In the merged image (F), the green and blue signals do not coincide. **G** and **H**, comparison of punctate staining in head [an enlargement of animal in (C)] with immunological staining of dense bodies with the monoclonal antibody MH35. Image (H) is reproduced from Waterston (1988). Anterior is to the left in all panels.

3.9 Bibliography

Chalfie, M,. Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. Science *263*, 802-805.

Fire, A., Harrison, S. W., and Dixon, D. (1990). A modular set of *lacZ* fusion vectors for studying gene expression in *Caenorhabditis elegans*. Gene *93*, 189-198.

Florio, S. K., Prusti, R. K., and Beavo, J. A. (1996). Solubilization of Membrane-bound Rod Phosphodiesterase by the Rod Phosphodiesterase Recombinant ∂ Subunit. J. Biol. Chem. *271*, 24036-24047.

Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. EMBO J. *10*, 3959-3970.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning, a laboratory manual*, 2nd. edition. Cold Spring Harbor Laboratory Press, NY.

Waterston, R. H. (1988). Muscle. in *The nematode* Caenorhabditis elegans. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

4 *C. elegans UNC-119* defines a new class of neural gene involved in axogenesis, and is functionally conserved with a human retinal protein

4.1 Introduction

A central goal in developmental neurobiology is to understand how individual axons navigate to produce the complex pattern of connectivity seen in the adult nervous system. Molecular identification of neural genes from diverse phyla has shown that pathfinding relies on homologous factors found in vertebrates and invertebrates, some of which were first identified in the nematode Caenorhabditis elegans (Tessier-Lavigne and Goodman, 1996). Its small size, nearly invariant cell lineage and established pattern of neuronal connectivity make *C. elegans* a tractable system in which to study neurogenesis (Brenner, 1974; Sulston and Horvitz, 1977; White *et al.*, 1986; Wood, 1988).

For example, the genes *unc-5*, *unc-6* and *unc-40* are involved in the migration of mesodermal cells and pioneering axons (Hedgecock, 1990). The UNC-6 protein is structurally related to the netrin family of guidance cues (Serafini *et al.*, 1994). Netrins establish dorsal-ventral gradients, providing cues to commissural interneurons in the developing vertebrate spinal cord (Kennedy *et al.*, 1994). Consistent with a conserved role, UNC-6 is expressed in pioneering neurons and neuroglia of the ventral midline of *C. elegans* (Wadsworth *et al.*, 1996). The *unc-40* gene, required primarily for ventral migrations (Hedgecock *et al.*, 1990), encodes a homologue of the DCC cell surface receptor (Chan *et al.*,

---

A version of this chapter has been submitted for publication. Maduro M and Pilgrim D (1998).

1996; Hedrick *et al.*, 1994; Fazeli *et al.*, 1997) which can bind netrin-1 (Keino-Masu *et al.*, 1996). Mutations of *Drosophila* and mouse DCC homologs also result in guidance defects (Kolodziej *et al.*, 1996; Fazeli *et al.*, 1997). The product of the *unc-5* gene, a novel cell surface receptor, is required for motile cells to be oriented away from ventral netrin sources (Hedgecock *et al.*, 1990; Leung-Hagesteijn *et al.*, 1992). Ectopic expression of *unc-5* in lateral sensory neurons causes their axons to grow dorsally instead of ventrally (Hamelin *et al.*, 1993), and experiments with vertebrate homologues of UNC-5 suggest that *unc-5* also encodes a netrin receptor (Leonardo *et al.*, 1997).

Compared to axon guidance, much less is understood about intrinsic factors affecting outgrowth, namely those that act in transducing guidance signals into growth cone motility (Tessier-Lavigne and Goodman, 1996). As many receptor families mediating outgrowth are conserved, it is likely that intracellular factors will be as well. For example, mutations in *C. elegans unc-33* result in abnormal axons with elevated numbers of microtubules (Hedgecock *et al.*, 1985; Li *et al.*, 1992). A chicken homolog of UNC-33, CRMP-62, is proposed to couple reception of external cues to growth cone collapse (Goshima *et al.*, 1995). Similarly, the UNC-76 protein, proposed to have an intracellular function in axon fasciculation, is structurally related to two human proteins, FEZ1 and FEZ2; FEZ1 can partially replace UNC-76 function in *C. elegans* (Bloom and Horvitz, 1997). The collection of such novel genes is expanding; owing to the complexity of growth cone motility, knowledge gained in all systems will likely be required for a complete understanding of outgrowth (Tanaka and Sabry, 1995; Tessier-Lavigne and Goodman, 1996). An important first step is the identification of axon outgrowth factors that are evolutionarily conserved, and hence likely to be important players.

We have previously reported the molecular identification of the novel *C. elegans unc-119* gene, and demonstrated a neural basis for the mutant phenotype (Maduro and Pilgrim, 1995). Mutants are defective in locomotion, pharyngeal pumping and dauer formation. Expression of *unc-119*::lacZ transgenes from both *C. elegans* and the related nematode *C. briggsae* occurs in most neurons throughout development, consistent with a neuron-specific function for UNC-119 (Maduro and Pilgrim, 1995, 1996). Shortly after we reported the sequence of *unc-119*, a similar open reading frame was identified by Higashide *et al.* (1996) in a screen for retina-specific cDNAs. The transcript of this gene (HRG4, for Human Retinal Gene 4), localizes specifically to rods and cones; expression begins at the onset of photoreceptor differentiation and is maintained through adulthood (Higashide *et al.*, 1996). The sequence similarity of UNC-119 and HRG4 suggest that they might be functional homologs.

In this paper, we extend the phenotypic analysis of *unc-119* and assess the function of HRG4 in *C. elegans*. First, it is shown indirectly that chemosensory neurons in *unc-119* animals have abnormal processes, and that the HSN motoneurons, which mediate egg laying behavior, have a presynaptic defect. Second, pan-neural GFP reporters are used to observe defective axons in the body wall and ventral nerve cord. We then test the ability of *unc-119* promoter-driven HRG4 expression to rescue the *unc-119* defects, and examine localization of functional UNC119::GFP and HRG4::GFP fusion transgenes in *unc-119* animals. Based on our results, we propose that UNC-119, HRG4 and DmUNC119, a newly identified *Drosophila* homolog, are members of a new class of proteins involved in axon guidance and fasciculation.

4.2 Materials and Methods

4.2.1 Construction of the unc-119::GFP nervous system marker

The clone pDP#MM016 contains the entire genomic *C. elegans unc-119* gene including 1.2 kb of putative control region (Maduro and Pilgrim, 1995). The plasmids pPD95.75 and pPD95.77 are GFP expression vectors (A. Fire, J. Ahnn, G. Seydoux and S. Xu, pers. comm.). The *unc-119*::GFP fusion plasmid pDP#MMUGF12 (abbreviated pUGF12; unc GFP fusion) was made by digesting pDP#MM016 with *Hin*dIII and *Pst*I and cloning into the same sites in pPD95.77. The resultant fusion gene encodes 101 amino acids of UNC-119 fused to the GFP coding region and 3'UTR from the *unc-54* gene and gives expression in the same cells as a similarly constructed *lacZ* fusion (Maduro and Pilgrim, 1995). No UNC-119 function is predicted to result from the amino-terminal region retained in pUGF12, as the null alleles *ed3* and *ed4* result from nonsense mutations corresponding to more downstream positions (Maduro and Pilgrim, 1995). The strong nervous system expression produced by transgenic arrays carrying pUGF12 does not appear to cause a phenotype of its own; no defects have been detected in N2 animals carrying such an array, and there is no evidence of rescue or alteration of the behavioral defects of transgenic *unc-119* mutants.

The functional *unc-119*::GFP fusion was constructed by introducing a novel *Bam*HI site into pDP#MM016 (to create pDP#MM016M) with primer MMA4 (see below) by site-directed mutagenesis (Kunkel, 1985). A *Bam*HI-*Eco*RI fragment containing the GFP cDNA was removed from pPD95.75 and cloned into the same sites in pUR#91 (a gift from W. Wadsworth). A slightly larger *Bam*HI-*Bgl*II fragment was then removed and cloned into the novel *Bam*HI site in pDP#MM016M, to create pUGF16.

4.2.2 Amplification of HRG4 and DmUNC119 cDNAs

The following oligonucleotides were used in RT-PCR (Sambrook *et al.*, 1989) and in the construction of HRG4 and DmUNC119 coding regions:

MMA14: (5') CCAGCCCAGAACTGGATCCTCCTGGG (3')
MMA19: (5') CGGTGGATCCGGG*ATG*GCGACGGAGTC (3')
MMA20: (5') TCAGGGGATCCCGCTGTAGGAATAGTC (3')
MMA32: (5') GCCGGATCCAAAATGCCGGGGCCGCTGCAGAG (3')
MMA33: (5') GCCGGATCCAAA*ATG*GTGACTCCCGACGAGGTG (3')
MMA34: (5') GCCGGATCCATATTGCCGCCATCGTAG (3')

The first-strand cDNA was synthesized from 1 µg of total human retina RNA (a gift from Ian MacDonald, University of Alberta) with primer MMA14. Mixed-stage total RNA was prepared with TRIzol reagent (BRL) from adults and larvae of the *Drosophila* strain *vg/CyO* (a gift from Ross Hodgetts, University of Alberta) and used for a first-strand reaction with primer MMA34. In both cases, Superscript II Reverse Transcriptase (BRL) was used with buffers supplied with the BRL 5'RACE kit. PCR was performed using MMA19/20 (for HRG4) or MMA33/34 (for DmUNC119), all of which generate *Bam*HI sites (underlined above), using the conditions (95°C, 5 min; 61°C, 45 sec; 72°C, 2 min) for one cycle, and (93°C, 1 min; 61°C, 45 sec; 73°C, 2 min) for 29 cycles using a Stratagene Robocycler 40 and reagents supplied with *Taq* polymerase (Amersham). The products were digested with *Bam*HI and cloned into pBluescript KS- (Stratagene) to generate pDP#MM092 (HRG4) and pDP#MM154 (DmUNC119). Both partial cDNAs lack an endogenous initiator codon; therefore, a novel ATG (shown italicized in the primer sequence) was introduced by each upstream primer. The cloning of the HRG4 cDNA was confirmed by dideoxy sequencing (Sanger *et al.*, 1977) using Sequenase 2.0 (Amersham) and cloning of DmUNC119 was confirmed using Thermo Sequenase (Amersham). The derivative HRG4 coding regions were obtained by PCR from pDP#MM103 using MMA32/20 for pDP#MM150 and MMA19/25 (see below) for pDP#MM148.

4.2.3 Construction of unc-119::HRG4 and unc-119::DmUNC119 fusions

Novel *Bam*HI sites were generated in pDP#MM016 immediately upstream of the endogenous stop (TAA) and start (ATG) codons using two rounds of oligonucleotide-mediated site-directed mutagenesis (Kunkel, 1985) with the following primers (*Bam*HI sites underlined, endogenous *unc-119* stop and start codons italicized):

MMA4: (5') CTACTCGCAGGATCCA*TAA*TTTCCCG (3')
MMA7: (5') GCTACAACAGGATCC*ATG*AAGGC (3')

The resultant plasmid was digested with *Bam*HI and the intervening segment containing the entire UNC-119 coding region was replaced with the various HRG4 and DmUNC119 inserts. Translation is predicted to initiate at the novel ATG of both fragments; the carboxy terminus is fused to the 3' noncoding region of *unc-119* that encodes an additional eight amino acids before a stop codon occurs in the case of HRG4, and one amino acid in the case of DmUNC119. The pDP#MM147 plasmid was obtained by *Pst*I digestion of pDP#MM103 and religation, which deletes an in-frame 42-bp fragment from the HRG4 coding region. The functional *unc-119*::HRG4::GFP fusion plasmid pUGF20 was made by first cloning the *Bam*HI-*Bgl*II fragment containing GFP (as described above for pUGF16) into pDP#MM016M, then placing the HRG4 cDNA into the recreated *Bam*HI site. The Genbank accession numbers for HRG4 and UNC-119 are U40998 (Maduro and Pilgrim, 1995) and U32854 (Higashide *et al.*, 1996) respectively.

4.2.4 Identification and cloning of Drosophila melanogaster homolog

Degenerate oligonucleotide pools were synthesized as shown below (*Bam*HI restriction sites are underlined; I = inosine; R = A or G; Y = C or T; W = A or T; S = G or C; M = A or C; H = A or C or T):

MMA23 (5') GACGGATCCGIATGATHGAIMGICAY (3')
MMA25 (5') CTGGGATCCACRAARTARAAISWRTC (3')

A single PCR product encoding part of a *D. melanogaster* homolog was amplified from 100 ng of Oregon R DNA with MMA23 and MMA25 using the following conditions: (95°C, 5 min; 53°C, 90 sec; 73°C, 5 min) for one cycle, and (93°C, 75 sec; 55°C, 75 sec; 73°C, 3 min) for 30 cycles. A single product containing part of HRG4 was also obtained from total human DNA under these conditions (data not shown). A larger genomic fragment of the *Drosophila* gene was obtained by screening a pBluescript KS- mini-library of ~5 kb genomic *HinD*III fragments, constructed using a procedure previously described for a size-selected *Ssf*I library of *C. elegans* (Maduro and Pilgrim, 1995). Coding regions were deduced by conservation with UNC-119/HRG4 and by prediction of splice sites, which were confirmed by RT-PCR (see above).

4.2.5 Construction of transgenic C. elegans strains

Microinjection of the syncytial gonad of young adult hermaphrodite worms was performed as described (Mello *et al.*, 1991). Plasmid DNA was prepared by alkaline lysis (Sambrook *et al.*, 1989) and injected at a concentration of 100 ng/µL. The presence of the transgenes was monitored by the Rol phenotype for the pRF4 marker, and rescue of the *unc-119* locomotory defect for the *C. elegans* clone pDP#MM016 (Maduro and Pilgrim, 1995), the *unc-119::*HRG4 fusion described above, and the functional GFP fusions.

To construct *unc-119(e2498)*; *edEx51*, the array (which contains pRF4 and pUGF12) was introduced using the following mating scheme: N2 males were crossed with *unc-119(e2498)*. The *unc/+* progeny males were mated into N2; *edEx51* hermaphrodites. Rol progeny were picked and self-fertilized to yield Unc Rol animals, which are homozygous for *unc-119(e2498)* and carry the *edEx51* extrachromosomal array.

Extrachromosomal arrays were integrated as follows. Fifty P_0 animals were treated with 1500 rad of γ irradiation from a ^{60}Co source, 300 transgenic F_1 animals were picked to self-fertilize individually, and F_2 animals were picked from plates showing >75% transmission. Homozygosity of the integrants was verified in the F_3 prior to backcrossing to N2.

4.2.6 Phenotypic assays

Analysis of dauer formation and quantitative analysis of locomotion, pharyngeal pumping, and egg retention were performed as described in Maduro and Pilgrim (1995). Dye-filling of the amphid neurons with fluorescein isothiocyanate (FITC, Sigma) was performed as described in Hedgecock *et al.* (1985). We performed pharmacological egg laying assays using the methods of Weinshenker *et al.* (1995), with serotonin (10 mg/mL, approximately 25 mM) or imipramine (0.75 mg/mL, approximately 2.4 mM) dissolved in M9 buffer (Wood, 1988).

4.2.7 Immunohistochemistry

DAPI (4', 6-diamino 2-phenylindole; Sigma) staining was performed by substituting 0.5 mg/mL DAPI for Xgal in the lacZ staining protocol described by Fire (1992). Animal fixation, permeabilization and antibody staining were performed as described (Desai *et al.*, 1988; Link *et al.*, 1992) using a rabbit anti-

GABA polyclonal antibody (1:100 dilution; Sigma) and a Cy3 conjugated secondary antibody (1:400 dilution, Jackson Immunochemicals). For detection of the HSN cell bodies, a polyclonal anti-serotonin antibody was used (1:100 dilution; INCSTAR). We were unable to detect the HSN processes with this antibody, however.

4.2.8 C. elegans microscopy

For GFP fluorescence, animals were mounted in 30 µL of 1% sodium azide (Sigma) in M9 buffer (Wood, 1988) or 0.1M NaCl on a microscope slide, covered with a coverslip and illuminated by a mercury lamp through the FITC filter set on a Zeiss Axioskop. For indirect immunofluorescence, 15 µL of a suspension of animals were added to 15 µL of DABCO (Sigma; powder dissolved at 25 mg/mL in 90% PBS-buffered glycerol, pH 8.6), covered with a coverslip, and viewed using the rhodamine filter set. Photographs were taken on Kodak Elite Slide Film (ASA 100) and scanned into digital form with a SprintScan 35 slide scanner. Adobe Photoshop was used to improve contrast and for annotation.

4.3 Results

4.3.1 unc-119 mutants have a presynaptic egg laying defect

We have previously described the behavioral phenotype of *unc-119* null mutants, that includes retention of eggs (Maduro and Pilgrim, 1995). In order to characterize the egg laying defect in *unc-119* animals, we used a pharmacological assay. Egg laying requires only a single pair of neurons, the HSNs, which innervate the egg laying muscles (White *et al.*, 1986; Desai and Horvitz, 1989). The HSNs cross-react to anti-serotonin antibodies, and exogenous serotonin causes release of eggs, suggesting that serotonin is released by the HSNs to stimulate contraction of the egg laying muscles (Desai *et al.*, 1988; Desai and Horvitz, 1989; Weinshenker *et al.*, 1995; Trent *et al.*, 1983). Exogenous imipramine can also stimulate egg laying, presumably by preventing serotonin reuptake (Trent *et al.*, 1983; Weinshenker *et al.*, 1995).

We tested the ability of *unc-119* mutants to respond to serotonin and imipramine (Figure 4.1D). In the absence of either drug, wild-type animals lay very few eggs over two hours ($\bar{x} = 0.6 \pm 0.3$), while *unc-119* animals show a significantly elevated level ($\bar{x} = 3.0 \pm 0.4$, $p < 0.0001$). This is consistent with our previous observation that *unc-119* mutants can lay eggs in the absence of food, which does not normally occur in the wild type (Maduro and Pilgrim, 1995). This suggests that regulation of egg laying behavior is disrupted. A similar defect has been suggested for *unc-34* and *unc-76* mutants, which have an egg laying phenotype similar to *unc-119* (Desai *et al.*, 1988).

As shown in Figure 4.1D, exposure to serotonin potentiates egg laying in *unc-119* mutants ($\bar{x} = 6.7 \pm 1.3$) to the same degree as in the wild type ($\bar{x} = 5.8 \pm 1.6$; $p > 0.50$) indicating that the postsynaptic response to the drug is normal (Trent *et al.*, 1983; Desai and Horvitz, 1989). In contrast, imipramine has no effect on egg laying in *unc-119* animals compared to controls ($\bar{x} = 2.3 \pm 0.4$; $p = 0.20$), while wild types demonstrate a significant increase ($\bar{x} = 5.5 \pm 1.0$; $p < 0.0002$). We are able to detect the presence of the HSN cell bodies in *unc-119* mutant animals with an anti-serotonin antibody (data not shown), suggesting that mutant HSNs have undergone at least partial differentiation (Desai *et al.*, 1988; Garriga and Stern, 1994) and that they make the neurotransmitter. The ability to respond to serotonin, but not imipramine, is characteristic of mutations affecting HSN structure or function (Trent *et al.*, 1983). We can infer that the egg laying phenotype of *unc-119* mutants results from a presynaptic defect, and not a defect in the egg laying muscles.

4.3.2 unc-119 mutants have abnormal chemosensory neurons

Exposure of *C. elegans* to the dye fluorescein isothiocyanate (FITC) fills six pairs of chemosensory neurons (amphids) in the head of the animal (Figure 4.3A), and two in the tail (phasmids), by direct entry through exposed cilia (Hedgecock *et al.*, 1985; Perkins *et al.*, 1986). While amphid cell bodies are observed to fill with dye in the wild type (Figure 4.3B), none of these fill in *unc-119(ed3)* mutants (Figure 4.3C). This is consistent with the inability of *unc-119* animals to form dauer larvae, as all other dye-filling mutants are also dauer-defective (Starich *et al.*, 1995). From epistasis tests, *unc-119* has been placed in the dauer formation pathway at a step that corresponds to mutations affecting cilium structure (Maduro and Pilgrim, 1995; Malone *et al.*, 1996). As all dye-filling mutants examined by electron microscopy show amphid structure defects (Perkins *et al.*, 1986; summarized by Starich *et al.*, 1995), we can predict that *unc-119* amphids are also abnormal. However, a defect in retrograde transport of the dye cannot be ruled out.

4.3.3 unc-119 mutants have multiple neuronal structure defects

The data thus far support a role for *unc-119* in neuron structure or function. To directly examine the possibilty of anatomical defects in *unc-119* mutants, the strong neural expression of the *unc-119* promoter itself was exploited (Maduro and Pilgrim, 1995). We constructed a fusion of *unc-119* to GFP, the Green Fluorescent Protein from *Aequoria victoria* (Chalfie *et al.*, 1994). Transgenic lines were obtained carrying the reporter as an extrachromosomal array (*edEx51*) or as chromosomal integrants (*edIs6, edIs7*; Mello *et al.*, 1991).

These strains accumulate cytoplasmic GFP in the same set of cells as a *lacZ* fusion described previously; expression begins in the pre-differentiation embryo in presumptive neurons and continues throughout adulthood in a large part of the nervous system (Maduro and Pilgrim, 1995). The *unc-119*::GFP marker therefore allows high-resolution visualization of neural processes in unfixed animals.

It is possible that the presence of the *unc-119*::GFP reporter, likely in multicopy arrays typical of transgenic DNA (Mello *et al.*, 1991), may itself affect neural structure, particularly since part of UNC-119 is present in the predicted fusion protein (see Materials and Methods). However, aside from the *rol-6D* transgene marker phenotype (Mello *et al.*, 1991), we observed no movement, development or other behavioral defects when these arrays are present in an otherwise wild-type background (data not shown). The gross morphology of the nervous system also appears completely normal in wild types (see below).

As a further control, we examined the ventral nerve cords and body wall axons in mutants known to contain abnormal processes. The *edIs6* marker was crossed into *sup-38*(*dv52*), *unc-33*(*e204*) and *unc-6*(*e78*), and mutants carrying the integrated *unc-119*::GFP transgene were examined by fluorescence microscopy (Figure 4.4). Consistent with previous reports, we observed abnormal axons in each case (Gilchrist and Moerman, 1992; S. Dames and C. Link, pers. comm.; Hedgecock *et al.*, 1985; Li *et al.*, 1992; McIntire *et al.*, 1992; Hedgecock *et al.*, 1990; Ishii *et al.*, 1992). We conclude that the *unc-119*::GFP marker is a reliable method for visualizing nervous system structure.

Transgenic *unc-119*; *edEx51* animals were examined by fluorescence microscopy, and found to contain two types of nervous system defects (Figures 4.5 and 4.6). First, all animals showed varying degrees of abnormal axon morphology, most visible in circumferential commissures and longitudinal processes. In the wild type (Figure 4.5B), commissures and longitudinal processes had a typical uniform, unbranched morphology (White *et al.*, 1986). In *unc-119* mutants, axons were often non-uniform, branched, and contained large varicosities (Figures 4.5C, 4.6A and 4.6B). Circumferential processes often demonstrated inappropriate turns consistent with defects in axonal guidance (McIntire *et al.*, 1992). Of the six classes of neurons required for *C. elegans* locomotion (VA, DA, VB, DB, VD, DD), all but two (VA and VB) have commissural axons (White *et al.*, 1986; White *et al.*, 1992). A structural defect in commissures is consistent with an "uncoordinated" phenotype (McIntire *et al.*, 1992), as seen in *unc-119* mutants.

The second defect observed was the defasciculation of the ventral nerve cord (VNC). The wild-type VNC consists of a large right process bundle, containing approximately forty axons, and a smaller left bundle, with only four axons (White *et al.*, 1986; Wightman *et al.*, 1997). The *unc-119*::GFP marker allows both bundles to be seen in wild type (Figure 4.6C). In *unc-119* mutants, the VNC consisted of abnormal longitudinal fascicles along its entire length (Figures 4.5C and 4.6D). The widths of these structures within the mutant VNC suggested that many axons were part of smaller process bundles consisting of transient fasciculations. Occasionally, we observed axons that emerged from the VNC and ran parallel to it for several hundred microns (Figure 4.6D).

To control for potential *unc-119*::GFP artifacts in *unc-119* mutants carrying the transgene, we examined axons using a reporter with a similar pan-neural expression pattern (F25B3.3::GFP; Bottorff, D., Ebinu, J., Chan, E., Stang, S., Dunn, R., Urban, S., Pilgrim, D., and Stone, J. C., submitted.). Similar branched neurites (Figure 4.6B) and defasciculated VNC were observed (Figure 4.6E). We also visualized GABAergic neurons using an anti-GABA antibody (McIntire *et al.*, 1993), and saw similar abnormalities (Figure 4.5E). We conclude that *unc-119* mutants have defects in axon guidance and fasciculation.

4.3.4 A functional UNC-119::GFP fusion is found in axons

A putative role in axon outgrowth, coupled with the nervous system expression of an *unc-119*::lacZ reporter (Maduro and Pilgrim, 1995), suggests that UNC-119 functions within axons. As a means of determining the possible intracellular localization of UNC-119, we constructed an *unc-119*::GFP fusion that retains the entire UNC-119 coding region, and therefore might also encode a functional UNC-119 fusion protein. The plasmid, pUGF16, contains the fusion of the *unc-119* genomic sequence to the GFP coding region, but retains the *unc-119* 3' flanking region (see Materials and Methods). Animals of genotype *unc-119*(*e2498*) were made transgenic for this plasmid by germline transformation (Mello *et al.*, 1991). Transgenic animals identified by neural GFP expression were found to have a wild-type phenotype, consistent with expression and function of the UNC-119::GFP fusion protein.

We obtained several independent chromosomal integrants for pUGF16 and examined the GFP localization in each strain. In all cases, the transgene expression pattern is similar to the original *unc-119*::lacZ fusion reported previously (Maduro and Pilgrim, 1995). Fluorescence appears within axons (whose morphology is normal due to rescue by the transgene), and is excluded from nuclei (Figure 4.7A). There is no evidence of subcellular localization; GFP signal is diffuse within neurons, although it appears stronger in cell bodies. With the caveat that expression pattern is not conclusive of functional necessity, these data are consistent with a role for UNC-119 within axons.

4.3.5 UNC-119 is predicted to have a human homolog based on sequence

The existence of a vertebrate homolog for UNC-119 might suggest an evolutionarily conserved role, as has been demonstrated for neural genes such as netrins (Kennedy *et al.*, 1994; Serafini *et al.*, 1994). At the time the predicted sequence of UNC-119 was reported, there were no similar proteins in the databases (Maduro and Pilgrim, 1995). Shortly thereafter, Higashide *et al.* (1996) identified a novel gene product, HRG4, expressed in the human retina, that we discovered to be similar to UNC-119. The similarities between the predicted amino acid sequences of HRG4 (240 amino acids) and UNC-119 (219 amino acids) are striking (Figure 4.8A). The proteins share strong similarity in two distinct regions. Region A shows 53% identity over 57 amino acids, and region B shows 67% identity over 112 residues (Figure 4.8B). The two proteins are dissimilar at their amino termini, where the majority of the additional amino acids in HRG4 appear, and in the space between regions A and B (15 amino acids in UNC-119). Despite this high degree of conservation, proteins with significant similarities in these domains have not been previously described. The similarities between UNC-119 and HRG4 have also been observed independently by another group (D. Zack, personal communication).

4.3.6 Identification of a novel Drosophila melanogaster homolog

The argument that UNC-119 is a conserved protein would be strengthened by the existence of a homolog in another divergent invertebrate system. We used the conservation between HRG4 and UNC-119 to construct a set of degenerate oligonucleotides for PCR (see Materials and Methods). A single, 444-bp product was obtained upon amplification from genomic DNA of the *Drosophila* melanogaster strain Oregon R. The PCR product was found to contain sequences similar to both HRG4 and UNC-119, and was mapped to the

X chromosome, band 7A1 (E. Woloshyn and D. Nash, unpublished results). A 5.0-kb *Hin*DIII genomic clone, pDP#MM121, was obtained from a plasmid mini-library of *Drosophila* DNA using the PCR product as a probe. A large portion of pDP#MM121 was sequenced and used to predict the sequence of the carboxyl 198 residues of the homolog, that we have named DmUNC119. The presence of two introns, predicted by comparison to HRG4, was confirmed by cloning and sequencing a partial cDNA obtained by RT-PCR.

An alignment with the nematode and human homologs demonstrates that the fly protein is more related to HRG4 than it is to UNC-119 (Figures 4.8A and 4.8B): DmUNC119 and HRG4 are 63% identical across region A, and 75% identical across region B. The intervening region between these two domains in DmUNC119 is the largest among the known UNC-119 homologs, and shares some additional amino acids in common with HRG4. To date, no *Drosophila* mutations have been correlated with a defect in DmUNC119.

4.3.7 The human homolog HRG4 can provide UNC-119 function in C. elegans

The strong sequence similarity shared by UNC-119, HRG4 and DmUNC119 suggests that these proteins may share the same biological activity. We examined whether the similarities in peptide sequence also extend to function by directly testing the ability of these gene products to replace UNC-119 in *C. elegans*. An *unc-119*::HRG4 fusion mini-gene, pDP#MM103, was constructed from a genomic DNA fragment containing *unc-119* in which the entire coding region was replaced by part of an HRG4 cDNA (Figure 4.9A). Nematodes homozygous for the *unc-119* null alleles *ed3* or *e2498* were made transgenic for pDP#MM103 by microinjection (Mello *et al.*, 1991). Among the progeny of injected mutants were animals that displayed a wild-type appearance, consistent with complementation of the *unc-119* mutation (Figure 4.2D).

A transmitting line, *unc-119*(*ed3*); *edEx49*, was obtained and used to quantify the degree of rescue of the behavioral defects. Locomotion, egg retention, pharyngeal pumping and dauer formation were restored by the *unc-119*::HRG4 transgene, in a manner indistinguishable from a wild-type *unc-119* genomic transgene (Figures 4.1A-4.1C, 4.2B, 4.2D). We further tested transgenic animals for rescue of the dye-filling and egg-laying response defects. Normal FITC uptake of amphid cell bodies was observed (Figure 4.3D), and egg-laying responses to serotonin and imipramine were not significantly different from wild type ($p > 0.50$; Figure 4.1D). Specifically, animals laid few eggs in buffer ($\bar{x} = 1.6 \pm 0.4$) and could be stimulated to lay eggs with imipramine ($\bar{x} = 7.7 \pm 1.6$; $p < 0.0001$).

We examined neuron morphology in mutants transgenic for *edEx50*, an array containing both the pUGF12 marker and the *unc-119*::HRG4 transgene pDP#MM103. As anticipated by the rescue of the behavioral defects, VNC and body wall axon morphology appeared normal (Figure 4.5D). Finally, anti-GABA immunohistochemistry revealed normal GABAergic neurons in *unc-119*(*e2498*); *edIs16*, a chromosomal integrant of the *unc-119*::HRG4 array *edEx49* (data not shown; McIntire *et al.*, 1993). We conclude from these results that HRG4 is a functionally conserved homolog of UNC-119.

4.3.8 A functional HRG4::GFP fusion is also found in axons

The ability of HRG4 to substitute for UNC-119 suggests that it must also be localized in a similar manner within neurons. As a means of following HRG4 expression, we inserted the same GFP coding region into the *unc-119*::HRG4 transgene as was used in the functional UNC-119::GFP fusion described above. The tripartite *unc-119*::HRG4::GFP fusion plasmid was also capable of rescuing the *unc-119* behavioral defects as an extrachromosomal array (data not shown).

We examined *unc-119*(*ed3*) animals, transgenic for this array, by fluorescence microscopy. The pattern of diffuse staining in axons was similar to the expression of the functional UNC-119::GFP fusion, although it was greatly reduced in intensity (Figure 4.7B). The accumulation of HRG4::GFP in axons suggests that it is functioning within neurons.

4.3.9 The carboxy terminus, but not the amino terminus, is required for HRG4 function in C. elegans

We have previously found that a genomic fragment of the *C. elegans* gene can confer partial rescue despite the absence of the promoter and coding region for the first 20 amino acids (Maduro and Pilgrim, 1995). It was of interest, therefore, to determine if the dissimilar HRG4 amino terminus is required for its *unc-119* rescuing activity. We tested altered versions of HRG4 for their ability to rescue the locomotory and nervous system defects of *unc-119* (Figure 4.9B).

Removal of the amino terminus of HRG4, up to the start of region A, did not affect the ability of HRG4 to rescue *unc-119* (plasmid pDP#MM150). Removal of a small part of the less-conserved region A (HRG4 amino acids 55-66) also had no effect, although this transgene still retained the amino terminus (pDP#MM147). In contrast, deletion of the carboxy terminal, which contains the longest contiguous region of identity among the three homologs (sequence VMHNKADY), completely abolished function (plasmid pDP#MM148). We conclude that in extrachromosomal transgenes, functional replacement of UNC-119 by HRG4 in *C. elegans* requires the carboxy terminus, but not the structurally divergent amino terminus.

4.3.10 The Drosophila homolog DmUNC119 can replace UNC-119 in C. elegans

Although the 5' end of the *Drosophila* mRNA and the coding region for the amino terminus have not been identified, the ability of an amino-truncated form of HRG4 to rescue the *unc-119* phenotype suggested that a similar portion of DmUNC119 might function in *C. elegans*. We made *unc-119(ed3)*; *him-8(e1489)* animals transgenic for an *unc-119*::DmUNC119 fusion plasmid, constructed in a manner similar to the amino-truncated HRG4 transgene. We found that these animals had normal locomotion and commissure structure, consistent with rescue of *unc-119* (Figure 4.9B). These preliminary results indicate that DmUNC119, like HRG4, is also a functionally conserved UNC-119 homolog.

4.4 Discussion

4.4.1 unc-119 is required for axon outgrowth and fasciculation

We have previously shown that *C. elegans unc-119* null mutants have multiple behavioral defects (Maduro and Pilgrim, 1995). In this work, we showed that mutants also have a presynaptic egg laying defect, and demonstrate weak uptake of the dye FITC into chemosensory neurons. A role in neural development, rather than function, is supported by the neuron structure defects observed in *unc-119* animals. First, commissures and other axons along the body wall showed abnormal branches, varicosities, and terminations. Second, ventral nerve cords were defasciculated. This panoply of defects suggests that *unc-119* belongs to a set of *C. elegans* genes required for proper axon outgrowth and fasciculation, which includes *unc-14*, *unc-33*, *unc-44*, *unc-51*, and *unc-73* (McIntire *et al.*, 1992). Indeed, the *unc-119* defects we have observed are very similar to those reported for the unrelated gene *unc-33*, which shows uncoordinated movement (Li *et al.*, 1992), a weak egg laying defect (Desai *et al.*, 1988), poor dye-filling with FITC (Hedgecock *et al.*, 1985), and defects in axon outgrowth and fasciculation (this work; McIntire *et al.*, 1992).

4.4.2 UNC-119, HRG4 and DmUNC119 define a new class of proteins

We observed the striking sequence similarity between UNC-119 and the human gene HRG4, and used the conservation to identify a new homolog, DmUNC119, from the fruit fly *Drosophila*. Both HRG4 and DmUNC119 are capable of replacing UNC-119 in *C. elegans*. Therefore, these proteins have been functionally conserved across metazoan evolution.

The existence of related genes from diverse phyla suggests that there are likely to be homologs in all metazoans. For example, there are known UNC-119 homologs in the related nematode *C. briggsae* (CbUNC119; Maduro and Pilgrim, 1996), the zebrafish *Danio rerio* (tentatively named DrUNC119; A. Manning and D.P., pers. comm.), and the rat *R. norvegicus* (RRG4; Higashide *et al.*, 1996). The UNC-119-related proteins define a new class, as they share no significant similarity with other known proteins. The only known exception is a predicted gene from the *C. elegans* genome sequencing project, C27H5.1 (Wilson *et al.*, 1994). This gene encodes a putative distant homolog of UNC-119 with limited amino acid similarity (Maduro and Pilgrim, 1996). Consistent with a potentially similar role to UNC-119, the C27H5.1 promoter drives GFP expression in many *C. elegans* neurons and in some cells of the pharynx (M. Chambers and D.P., unpublished results). Given the existence of a second *unc-119*-like gene within *C. elegans*, it is likely that there are other distant homologs to be found in other systems.

4.4.3 Possible UNC-119/HRG4 functions

The ability of HRG4 and DmUNC119 to compensate for loss of *unc-119* function in *C. elegans* demonstrates that these proteins have similar function, and that they likely interact with other conserved molecules. The phenotype of *unc-119* mutants, and the expression of transgene fusions, suggest that HRG4 may have a role in the vertebrate nervous system. However, the expression of HRG4 appears to be strongly enriched in retina, as mRNA from other tissues does not contain detectable amounts of the message (Higashide *et al.*, 1996). Localization of the HRG4 mRNA to rod and cone photoreceptors raised the possibility that HRG4 might function in photoreception (Higashide *et al.*, 1996). Although a weak photoresponse has been proposed in *C. elegans* (Burr, 1985), it is unlikely that *unc-119* contributes to light perception because of its implied function in both motor neurons and sensory neurons. Conversely, it is not known if HRG4 could function in neurite outgrowth: although axon guidance has been extensively studied in ganglionic cells (e.g. Stier and Schlosshauer, 1995), the existence of a growth cone mechanism for rod and cone cell development has not been established.

Without further biochemical or genetic data, is difficult to assign a precise role to the UNC-119/HRG4 class of proteins. The absence of known structural or signal motifs in the UNC-119 and HRG4 amino acid sequences is consistent with an intracellular role, which is further supported by the ability of functional transgene fusions to target GFP to axons. As anticipated by the abnormal structure of mutant axons and fascicles, UNC-119 could participate in coordinating cytoskeletal changes required for outgrowth and adhesion with external growth cues (Tanaka and Sabry, 1995). We postulate, therefore, that UNC-119/HRG4 might have a direct role in establishment of proper cytoskeletal structure, or it may have a more indirect role in signal transduction. There are precedents for an assignment to either function from within the set of *C. elegans* genes having multiple effects on axogenesis similar to *unc-119*: Mutant *unc-33* chemosensory axons contain a superabundance of microtubules, consistent with a cytoskeletal defect (Hedgecock *et al.*, 1985); and the *unc-51* gene encodes a putative serine/threonine kinase (Ogura *et al.*, 1994).

A further understanding of UNC-119/HRG4 function awaits the determination of interacting proteins and determination of *in vivo* cellular localization. Insight may come from genetic analysis in other systems, particularly in vertebrates, as there are no documented human diseases known to be associated with mutation in HRG4.

Acknowledgements

We would like to thank Joel Rothman, Gian Garriga, Jeff Goldberg, and Bob Campenot for critical review of the original manuscript, and for providing helpful suggestions. We would also like to thank the Caenorhabditis Genetics Center for providing some of the strains used in this work; Chris Link for donating *sup-38*(*dv52*), for providing invaluable help and materials for immunohistochemistry, and for general advice; Andy Fire for providing GFP expression vectors; Jeff Goldberg for antibodies; Dave Weinshenker for extensive help with the imipramine assays; Ian Macdonald for the total human retina RNA; Gary Ritzel for the *Drosophila* DNA; Effie Woloshyn and David Nash for *in-situ* mapping of DmUNC119; Chris Link, Shale Dames, Angela Manning, Michelle Chambers, Sinisa Urban, Don Zack for contributing unpublished results and materials; Paul Wong for advice; Jiangwen Zhu for protocols on integrating extrachromosomal arrays; Gina Broitman for sequencing the DmUNC119 partial cDNA clone; and Bruce Draper for initially suggesting the use of *unc-119*::GFP as a nervous

system marker in *unc-119* mutants. This work was supported by grants from the Medical Research Council of Canada and the Natural Sciences and Engineering Research Council.

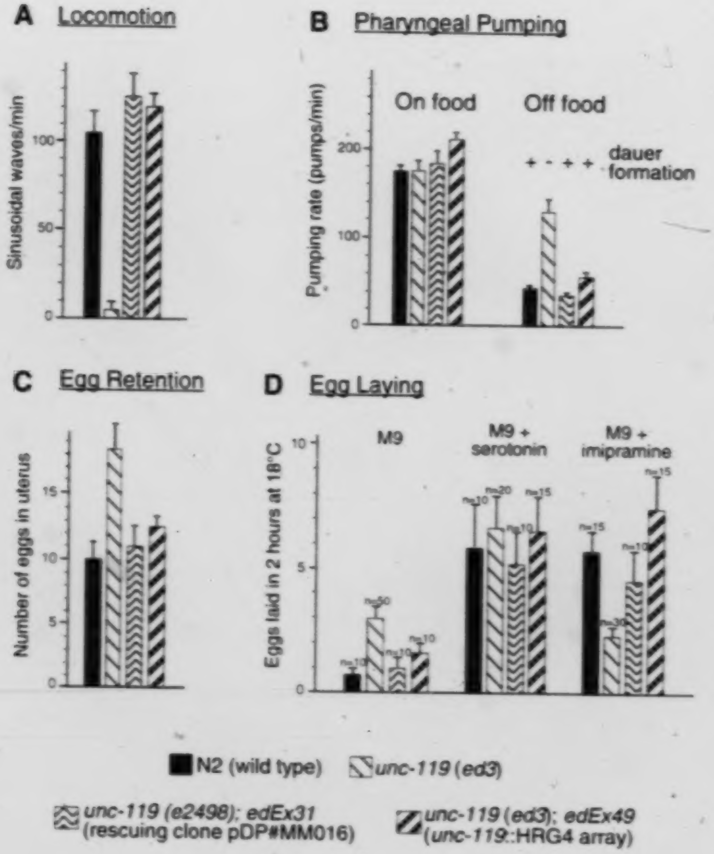


Figure 4.1 Quantification of *unc-119* behavioral defects and rescue by transgenes. At least ten animals were scored for each strait. Data collection for panels (A)-(C) is described in Maduro and Pilgrim (1995). Data collection for (D) is described in the text. Lines above data bars denote standard error of the mean (SEM). **A**, Locomotion measured in liquid (M9 buffer). **B**, Pharyngeal pumping. '+' denotes ability to form dauer larvae, and '-' denotes a dauer formation defect. **C**, Retention of eggs in adult hermaphrodites in the presence of food. **D**, Potentiation of egg laying in the presence of serotonin (10 mg/mL, approximately 25 mM) or imipramine (0.75 mg/mL, approximately 2.4 mM).

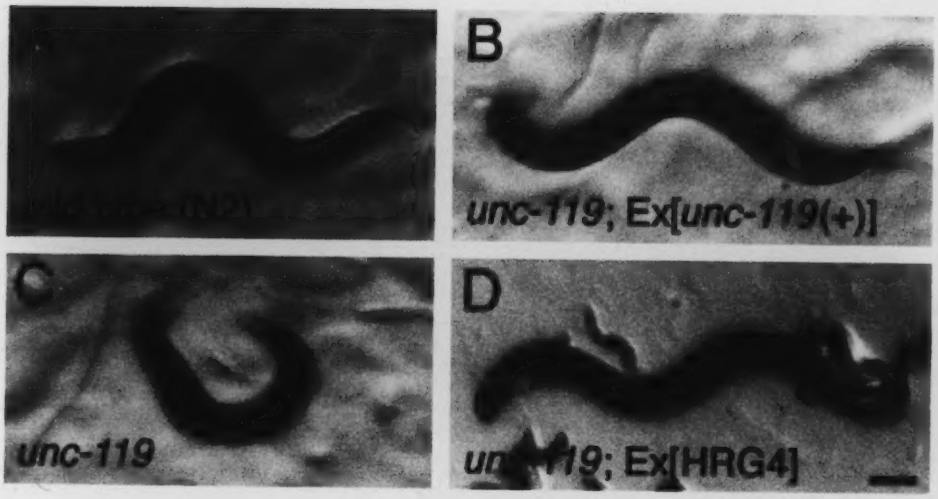


Figure 4.2 Still frames from video images of live animals on agar plates. Anterior is shown to the left. Scale bar, 100 *µ*m. ***A***, Wild type nematode showing typical dorsal-ventral sinusoidal waves propagated during motion. ***B***, *unc-119*(*e2498*) carrying the extrachromosomal array *edEx31*, which carries a plasmid copy of the *unc-119* gene. Note normal sinusoidal wave. ***C***, *unc-119*(*ed9*) mutant displaying ventral coiling and paralysis. ***D***, *unc-119*(*ed3*) carrying the array *edEx51*, which contains the *unc-119*::HRG4 fusion. Note normal wave as in ***A*** and ***B***.

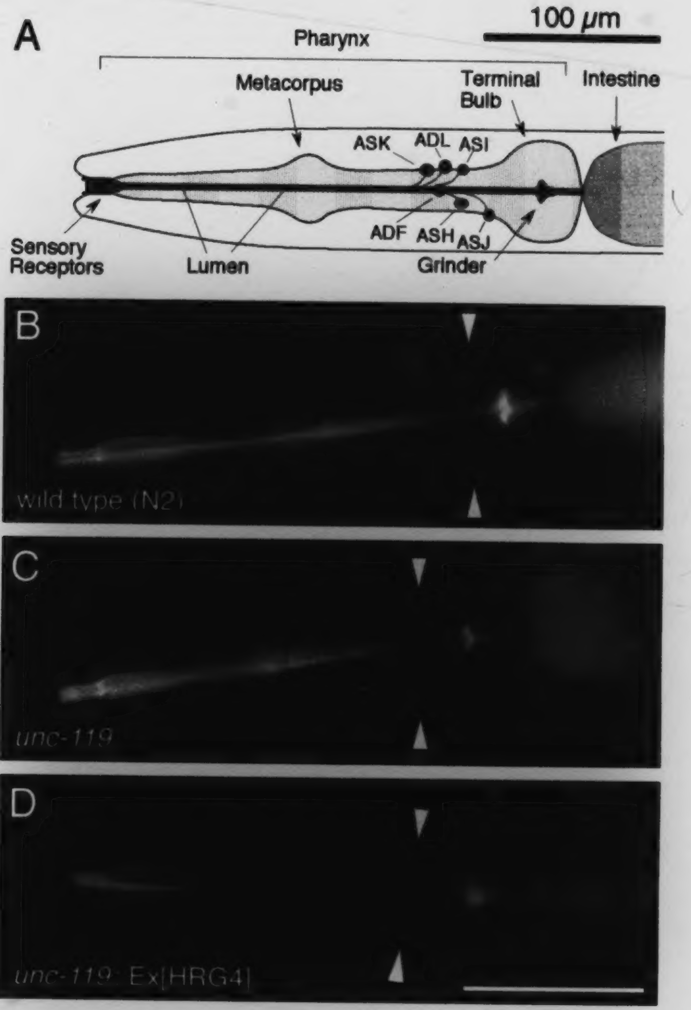


Figure 4.3 Uptake of the fluorescent dye FITC into amphid sensory neurons. All views shown are lateral with anterior to the left. Scale bar, 100 μm.
A, Diagram of head of animal showing the structure of the pharynx and the positions cell bodies that fill with FITC. Figure is adapted from Starich *et al.* (1995) and Wadsworth and Hedgecock (1996). ***B***, Wild type (N2) animal demonstrates dye filling (arrowheads). ***C***, In an *unc-119*(*ed3*) animal no dye filling is observed (arrowheads). ***D***, *unc-119*(*ed3*) animal carrying *unc-119*::HRG4 transgene demonstrates rescue of the dye filling defect. Other cell bodies may be out of the plane of focus. Note that in (B)-(D), additional fluorescence comes from entry of the dye into the pharyngeal lumen, and from fluorescence of gut granules in the intestine.

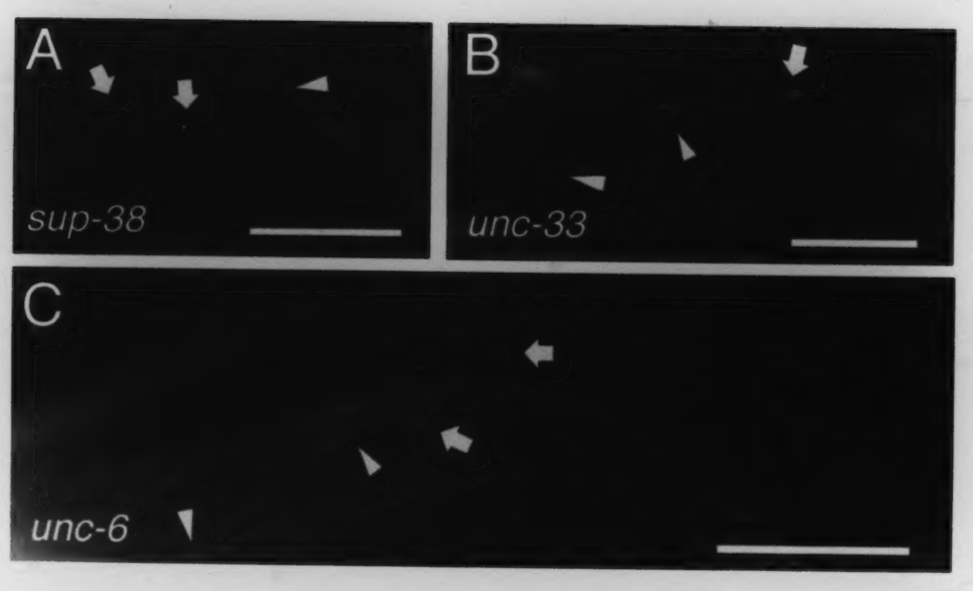


Figure 4.4 GFP fluorescence in neurons of three unrelated genes known to have axon defects. In each case, varicosities (arrows) and crossed fascicles can be seen (arrowheads). All views shown are lateral with anterior to the left. Scale bar, 50 μm. (A-C) All animals carry *edIs6*, a chromosomal integration of *unc-119*::GFP and pRF4 (*rol-6D*). ***A***, *sup-38*(*dv52*); *edIs6*. ***B***, *unc-33*(*e204*); *edIs6*. ***C***, *unc-6*(*e78*); *edIs6*.

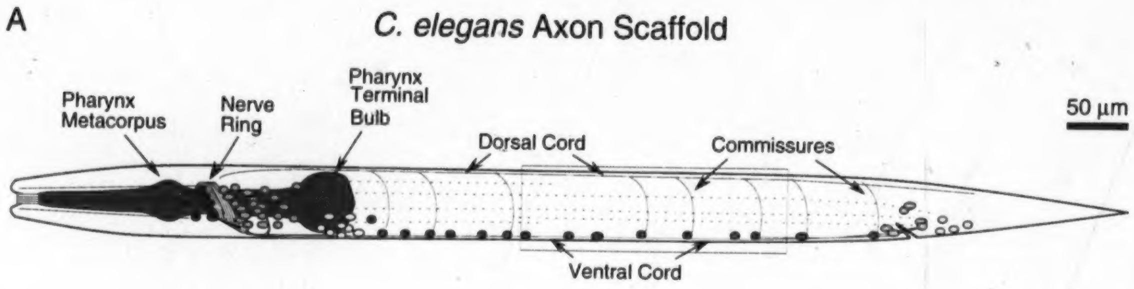


Figure 4.5 Axon morphology in wild type, *unc-119* and rescued animals. Anterior is shown to the left. Yellow punctate fluorescence is due to gut granules, not GFP expression. Scale bars, 50 µm.
A, Diagram of major neuron processes in C. elegans. The region of animals shown in panels (B)-(E) is indicated by a rectangle. (Continued on next page)

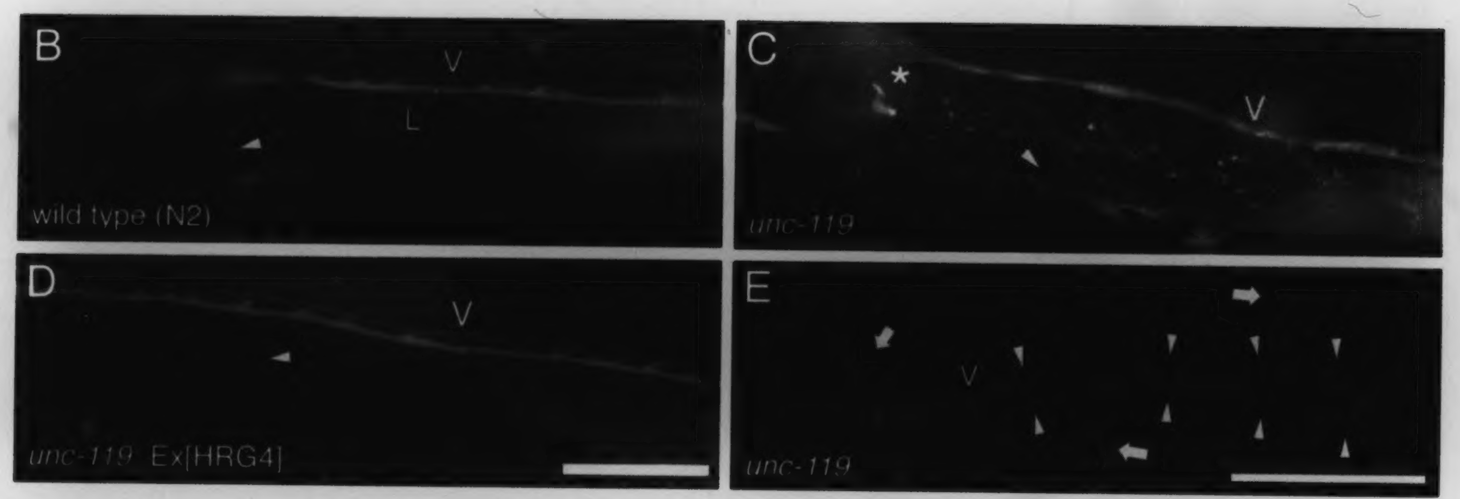


Figure 4.5, continued. **B**, Subventral view of N2; *edEx51* animal showing normal ventral nerve cord (V), longitudinal processes (L) and commissural axons (e.g. arrowhead). **C**, Subventral view of *unc-119*(*e2498*); *edEx51* animal shows abnormal ventral nerve cord (V) and abnormal body wall neurons (arrowhead). The bifurcation of the ventral nerve cord around the vulva is shown with an asterisk (*). **D**, *unc-119*(*ed3*); *edEx50* animals (rescued by the *unc-119*::HRG4 transgene). The ventral nerve cord (V) and commissures (e.g. arrowhead) have normal morphology, similar to panel (B). **E**, Anti-GABA staining of *unc-119*(*e2498*) animal showing abnormal processes (arrowheads) in the ventral nerve cord (V), and crossed or looped axons (arrows). The extrachromosomal array *edEx51* contains the *unc-119*::GFP fusion pUGF12 and the marker plasmid pRF4 (*rol-6D*). The array *edEx50* contains pUGF12 and the *unc-119*::HRG4 fusion.

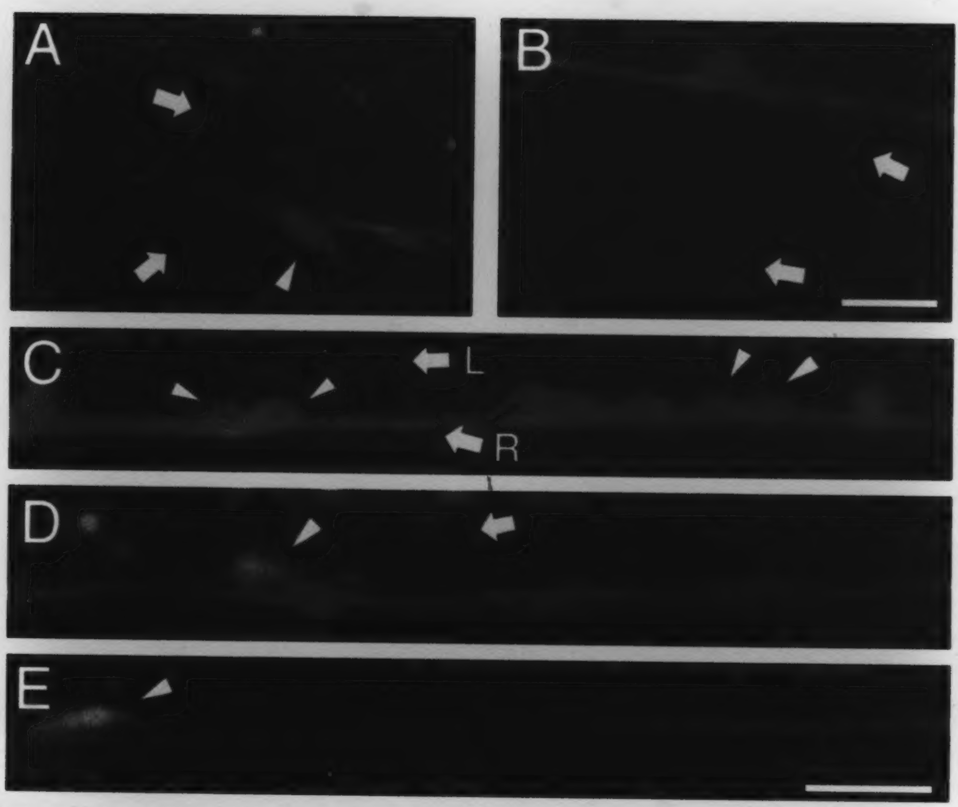


Figure 4.6 Detailed views of *unc-119* axon defects. Panels (A) and (B) are at similar magnification, as are panels (C) through (E). Anterior is to the left in all panels. Views in (A) and (B) are lateral, while those in (C)-(E) are ventral. Yellow fluorescence is due to gut granules. Scale bars, 10 μm. **A**, Branched axons (arrows) and large varicosity (arrowhead) in *unc-119*(*e2498*); *edEx51* (higher magnification of animal in Figure 4.5C). **B**, *unc-119*(*e2498*) animal carrying an F25B3.3::GFP reporter as an extrachromosomal array. Short commissural axons emerge from a longitudinal process bundle, terminating in branches (arrows). **C**, Ventral nerve cord in wild type carrying *edEx51*. The main right process bundle (R) and smaller left bundle (L) are indicated. Cell bodies (arrowheads) are visible adjacent to the neuropile. **D**, *unc-119*(*e2498*); *edIs6* animal shows that processes still run along the ventral midline, but are defasciculated. Occasionally, a few processes run parallel to the defasciculated bundle (arrow). **E**, *unc-119*(*e2498*) animal carrying F25B3.3::GFP also shows abnormal ventral nerve cord.

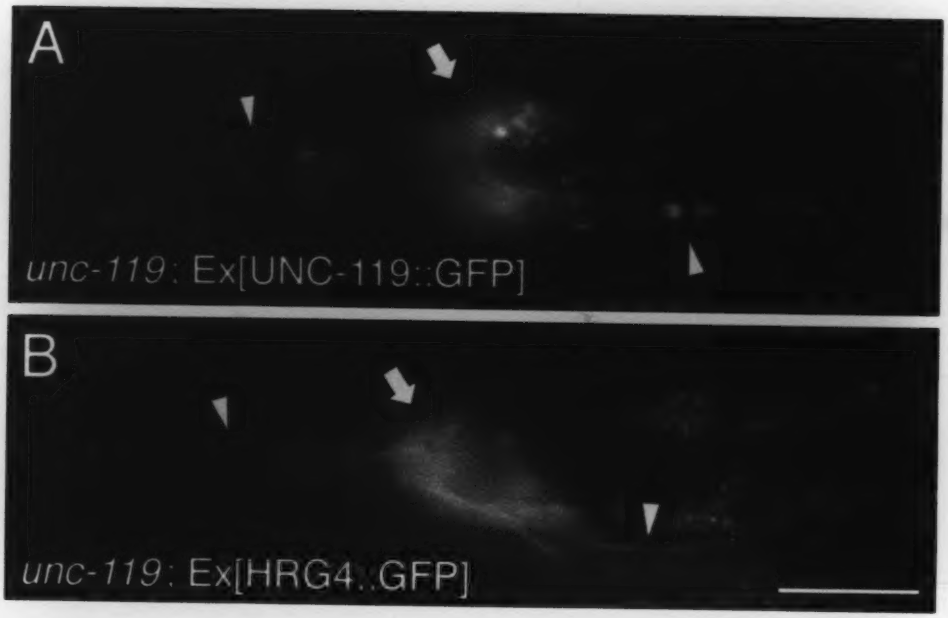


Figure 4.7 Functional GFP fusions are found in axons. The strong fluorescence of the nerve ring is indicated by an arrow in each panel (compare Figure 5A). Anterior is shown to the left. Scale bar, 50 μm. **A**, *unc-119*(*ed9*) animal homozygous for *edIs8*, a chromosomal integrant of a functional *unc-119*::UNC-119::GFP fusion (pUGF16). **B**, *unc-119*(*ed3*) animal carrying an extrachromosomal *unc-119*::HRG4::GFP fusion (pUGF20). In both panels, GFP is excluded from neurons and found in axons (arrowheads). This image was overexposed to reveal the weaker fluorescence demonstrated by the pUGF20 transgene; this also enhances appearance of gut granules (punctate yellow fluorescence).

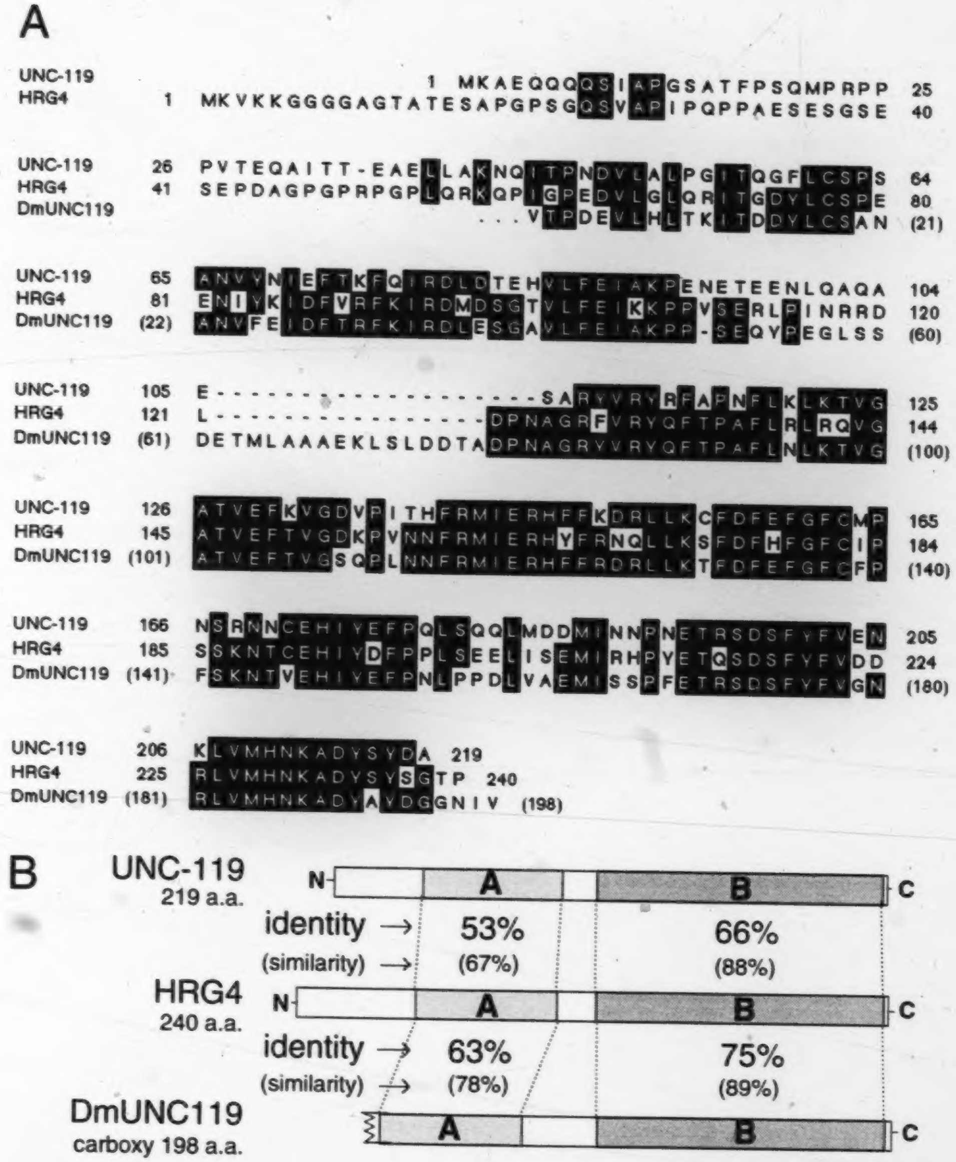

A
UNC-119 1 MKAEQQQQSIAPGSATFPSQMPRPP 25
HRG4 1 MKVKKGGGGAGTATESAPGPSGQSVAPIPQPPAESESGSE 40
UNC-119 26 PVTEQAITT-EAELLAKNQITPNDVLALPGITQGFLCSPS 64
HRG4 41 SEPDAGPGPRPGPLQRKQPIGPEDVLGLQRITGDYLCSPE 80
DmUNC119 ...VTPDEVLHLTKITDDYLCSAN (21)
UNC-119 65 ANVYNIEFTKFQIRDLDTEHVLFEIAKPENETEENLQAQA 104
HRG4 81 ENIYKIDFVRFKIRDMDSGTVLFEIKKPPVSERLPINRRD 120
DmUNC119 (22) ANVFEIDFTRFKIRDLESGAVLFEIAKPP-SEQYPEGLSS (60)
UNC-119 105 E---------------------SARYVRYRFAPNFLKLKTVG 125
HRG4 121 L-----------------DPNAGRFVRYQFTPAFLRLRQVG 144
DmUNC119 (61) DETMLAAAEKLSLDDTADPNAGRYVRYQFTPAFLNLKTVG (100)
UNC-119 126 ATVEFKVGDVPITHFRMIERHFFKDRLLKCFDFEFGFCMP 165
HRG4 145 ATVEFTVGDKPVNNFRMIERHYFRNQLLKSFDFHFGFCIP 184
DmUNC119 (101) ATVEFTVGSQPLNNFRMIERHFFRDRLLKTFDFEFGFCFP (140)
UNC-119 166 NSRNNCEHIYEFPQLSQQLMDDMINNPNETRSDSFYFVEN 205
HRG4 185 SSKNTCEHIYDFPPLSEELISEMIRHPYETQSDSFYFVDD 224
DmUNC119 (141) FSKNTVEHIYEFPNLPPDLVAEMISSPFETRSDSFYFVGN (180)
UNC-119 206 KLVMHNKADYSYDA 219
HRG4 225 RLVMHNKADYSYSGTP 240
DmUNC119 (181) RLVMHNKADYAVDGGNIV (198)
B
UNC-119
219 a.a.
N
A
B
C
identity →
(similarity) →
53%
(67%)
66%
(88%)
HRG4
240 a.a.
N
A
B
C
identity →
(similarity) →
63%
(78%)
75%
(89%)
DmUNC119
carboxy 198 a.a.
A
B
C

Figure 4.8 Comparison of *C. elegans* UNC-119, human HRG4 and partial *Drosophila* DmUNC119. **A**, Amino acid alignment. Identities present in at least two of the three genes are highlighted. The amino acid positions of DmUNC119 are indicated starting from the valine corresponding to HRG4 position 60. **B**, Schematic alignment of the three homologs showing the two regions A and B, and the percentage identities and conserved changes (similarities) with HRG4.

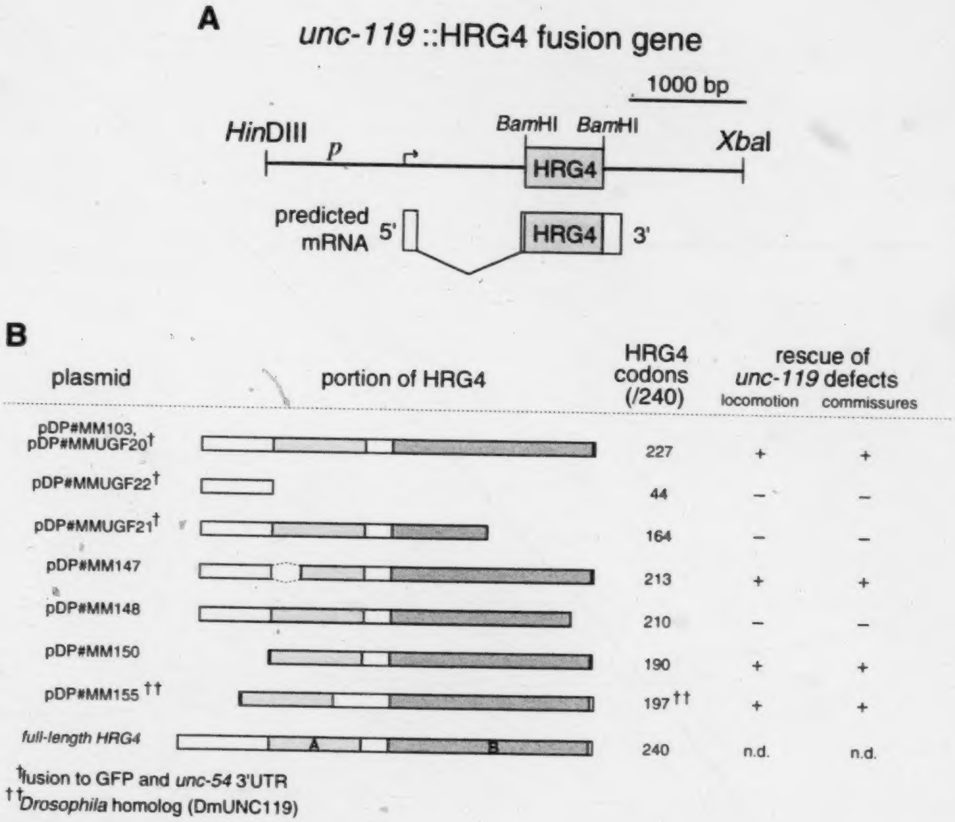


Figure 4.9 Rescuing ability of HRG4 and DmUNC119 transgenes. **A**, Diagram of the *unc-119*::HRG4 fusion gene used to test for rescue by the human homolog, and its predicted mRNA. **B**, Rescue of locomotory and commissural defects by various HRG4 and DmUNC119 transgenes. The structure of the predicted full-length HRG4 protein is shown for comparison; conserved regions A and B are shaded. 'HRG4 codons' indicates the number of amino acids of HRG4 contained in each transgene. At least ten transgenic animals were examined in each case for rescue of the locomotory and commissure defects. Plasmids pDP#MM103, 147, 148 and 150 were injected into *unc-119*(*e2498*), the GFP fusions pUGF20-22 were injected into *unc-119*(*ed3*), and pDP#MM155 was injected into *unc-119*(*ed3*); *him-8*(*e1489*). Plasmids pDP#MM147-155 were coinjected with pRF4 (*rol-6*) and pUGF12 (*unc-119*::GFP) as transgenic markers and to allow visualization of commissures.

4.5 Bibliography

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. NAR *25*, 3389-3402.

Bloom, L., and Horvitz, H. R. (1997). The Caenorhabditis elegans gene *unc-76* and its human homologs define a new gene family involved in axonal outgrowth and fasciculation. PNAS *94*, 3414-3419.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77*, 71-94.

Burr, A. H. (1985). The photomovement of Caenorhabditis elegans, a nematode which lacks ocelli. Proof that the response is to light not radiant heating. Photochem. Photobiol. *41*, 577-582.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. Science *263*, 802-805.

Chan, S. S.-Y., Zheng, H., Su, M.-W., Wilk, R., Killeen, M. T., Hedgecock, E. M., and Culotti, J. G. (1996). UNC-40, a *C. elegans* Homolog of DCC (Deleted in Colorectal Cancer), Is Required in Motile Cells Responding to UNC-6 Netrin Cues. Cell *87*, 187-195.

Desai, C., Garriga, G., McIntire, S. L., and Horvitz, H. R. (1988). A genetic pathway for the development of the Caenorhabditis elegans HSN motor neurons. Nature *336*, 638-646.

Desai, C., and Horvitz, H. R. (1989). Caenorhabditis elegans mutants defective in the functioning of the motor neurons responsible for egg laying. Genetics *121*, 703-721.

Fazeli, A., Dickinson, S. L., Hermiston, M. L., Tighe, R. V., Steen, R. G., Small, C. G., Stoeckli, E. T., Keino-Masu, K., Masu, M., Rayburn, H., Simons, J., Bronson, R. T., Gordon, J. I., Tessier-Lavigne, M., and Weinberg, R. A. (1997). Phenotype of mice lacking functional Deleted in colorectal cancer (Dcc) gene. Nature *386*, 796-804.

Fire, A. (1992). Histochemical techniques for locating Escherichia coli β-galactosidade activity in transgenic organisms. Genet. Anal. Tech. Appl. *9*, 152-160.

Garriga, G., and Stern, M. J. (1994). Hams and Egls: Genetic analysis of cell migration in Caenorhabditis elegans. Curr. Opin. Genet. Dev. *4*, 575-580.

Gilchrist, E. J., and Moerman, D. G. (1992). Mutations in the *sup-38* gene of Caenorhabditis elegans suppress muscle-attachment defects in *unc-52* mutants. Genetics *132*, 431-442.

Goshima, Y., Nakamura, F., Strittmatter, P., and Streittmatter, S. M. (1995). Collapsin-induced growth cone collapse mediated by an intracellular protein related to UNC-33. Nature *376*, 509-514.

Hamelin, M., Zhou, Y., Su, M.-W., Scott, I. M., and Culotti, J. G. (1993). Expression of the UNC-5 guidance receptor in the touch neurons of *C. elegans* steers their axons dorsally. Nature *364*, 327-330.

Harris, R., Sabatelli, L. M., and Seeger, M. A. (1996). Guidance Cues at the *Drosophila* CNS Midline: Identification and Characterization of Two *Drosophila* Netrin/UNC-6 Homologs. Neuron *17*, 217-228.

Hedgecock, E. M., Culotti, J. G., and Hall, D. H. (1990). The *unc-5*, *unc-6*, and *unc-40* genes guide circumferential migrations of pioneer axons and mesodermal cells on the epidermis in *C. elegans*. Neuron *2*, 61-85.

Hedgecock, E. M., Culotti, J. G., Thomson, J. N., and Perkins, L. A. (1985). Axonal guidance mutants of Caenorhabditis elegans identified by filling sensory neurons with fluorescein dyes. Dev. Biol. *111*, 158-170.

Hedrick, L., Cho, R. R., Fearon, E. R., Wu, T. C., Kinzler, K. W., and Vogelstein, B. (1994). The DCC gene product in cellular differentiation and colorectal tumorigenesis. Genes Dev. *8*, 1174-1183.

Higashide, T., Murakami, A., McLaren, M. J., and Inana, G. (1996). Cloning of the cDNA for a Novel Photoreceptor Protein. J. Biol. Chem. *271*, 1797-1804.

Ishii, N., Wadsworth, W. G., Stern, B. D., Culotti, J. G., and Hedgecock, E. M. (1992). UNC-6, a Laminin-Related Protein, Guides Cell and Pioneer Axon Migrations in *C. elegans*. Neuron *9*, 873-881.

Keino-Masu, K., Masu, M., Hinck, L., Leonardo, E. D., Chan, S. S.-Y., Culotti, J. G., and Tessier-Lavigne, M. (1996). Deleted in Colorectal Cancer (DCC) encodes a netrin receptor. Cell *87*, 175-185.

Kennedy, T. E., Serafini, T., de la Torre, J. R., and Tessier-Lavigne, M. (1994). Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord. Cell *78*, 425-435.

Kolodziej, P. A., Timpe, L. C., Mitchell, K. J., Fried, S. R., Goodman, C. S., Jan, L. Y., and Jan, Y. N. (1996). frazzled encodes a *Drosophila* member of the DCC immunoglobulin subfamily and is required for CNS and motor axon guidance. Cell *87*, 197-204.

Komatsu, H., Mori, I., Rhee, J.-S., Akaike, N., and Ohshima, Y. (1996). Mutations in a Cyclic Nucleotide-Gated Channel Lead to Abnormal Thermosensation and Chemosensation in *C. elegans*. Neuron *17*, 707-718.

Kunkel, T. A. (1985). Rapid and specific site-specific mutagenesis without phenotypic selection. PNAS *82*, 488-492.

Leonardo, E. D., Hinck, L., Masu, M., Keino-Masu, K., Ackerman, S. L., Tessier-Lavigne, M. (1997) Vertebrate homologues of *C. elegans* UNC-5 are candidate netrin receptors. Nature *386*, 833-838.

Lewis, J. A., Wu, C. H., Berg, H., and Levine, J. H. (1980). The genetics of levamisole resistance in the nematode Caenorhabditis elegans. Genetics *95*, 905-928.

Leung-Hagesteijn, C., Spence, A. M., Stern, B. D., Zhou, Y., Su, M.-W., Hedgecock, E. M., and Culotti, J. G. (1992). UNC-5, a Transmembrane Protein with Immunoglobulin and Thrombospondin Type 1 Domains, Guides Cell and Pioneer Axon Migrations in *C. elegans*. Cell *71*, 289-299.

Li, W., Herman, R. K., and Shaw, J. E. (1992). Analysis of the Caenorhabditis elegans axonal guidance and outgrowth gene *unc-33*. Genetics *132*, 675-689.

Link, C. D., Silverman, M. A., Breen, M., and Watt, K. E. (1992). Characterisation of Caenorhabditis elegans lectin-binding mutants. Genetics *131*, 867-881.

Maduro, M., and Pilgrim, D. (1995). Identification and Cloning of *unc-119*, a Gene Expressed in the Caenorhabditis elegans Nervous System. Genetics *141*, 977-988.

Maduro, M., and Pilgrim, D. (1996). Conservation of function and expression of *unc-119* from two Caenorhabditis species despite divergence of non-coding DNA. Gene *183*, 77-85.

Malone, E. A., Inoue, T., and Thomas, J. H. (1996). Genetic Analysis of the Roles of *daf-28* and *age-1* in Regulating Caenorhabditis elegans Dauer Formation. Genetics *143*, 1193-1205.

McIntire, S. L., Garriga, G., White, J., Jacobson, D., and Horvitz, H. R. (1992). Genes necessary for directed axonal elongation or fasciculation in *C. elegans*. Neuron *8*, 307-322.

McIntire, S. L., Jorgensen, E., Kaplan, J., and Horvitz, H. R. (1993). The GABAergic nervous system of Caenorhabditis elegans. Nature *364*, 337-341.

Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. EMBO J. *10*, 3959-3970.

Ogura, K., Wicky, C., Magnenat, L, Tobler, H., Mori, I., Muller, F., and Ohshima, Y. (1994). Caenorhabditis elegans *unc-51* gene required for axonal elongation encodes a novel serine/threonine kinase. Genes Dev *84*, 2389-2400.

Perkins, L. A., Hedgecock, E. M., Thomson, J. N., and Culotti, J. G. (1986). Mutant sensory cilia in the nematode Caenorhabditis elegans. Dev. Biol. *117*, 456-487.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. PNAS *74*, 5463-5467.

Serafini, T., Kennedy, T. E., Galko, M. J., Mirzayan, C., Jessell, T. M., and Tessier-Lavigne, M. (1994). The netrins define a family of axon outgrowth-promoting proteins homologous to *C. elegans* UNC-6. Cell *78*, 409-424.

Starich, T. A., Herman, R. K., Kari, C. K., Yeh, W.-H., Schackwitz, W. S., Schuyler, M. W., Collet, J., Thomas, J. H., and Riddle, D. L. (1995). Mutations Affecting the Chemosensory Neurons of Caenorhabditis elegans. Genetics *139*, 171-188.

Stier, H., and Schlosshauer, B. (1995). Axonal guidance in the chicken retina. Development *121*, 1443-1454.

Sulston, J. E., and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode Caenorhabditis elegans. Developmental Biology *56*, 110-156.

Tanaka, E., and Sabry, J. (1995). Making the Connection: Cytoskeletal Rearrangements during Growth Cone Guidance. Cell *83*, 171-178.

Tessier-Lavigne, M., and Goodman, C. S. (1996). The Molecular Biology of Axon Guidance. Science *274*, 1123-1133.

Trent, C., Tsung, N., and Horvitz, H. R. (1983). Egg-laying defective mutants of the nematode Caenorhabditis elegans. Genetics *104*, 619-647.

Wadsworth, W. G., Bhatt, H., and Hedgecock, E. M. (1996). Neuroglia and Pioneer Neurons Express UNC-6 to Provide Global and Local Netrin Cues for Guiding Migrations in *C. elegans*. Neuron *16*, 35-46.

Weinshenker, D., Garriga, G., and Thomas, J. H. (1995). Genetic and pharmacological analysis of neurotransmitters controlling egg laying in *C. elegans*. J. Neurosci. *15*, 6975-6985.

White, J.G., Southgate, E., and Thomson, J. N. (1991). Mutations in the Caenorhabditis elegans *unc-4* gene alter the synaptic input to ventral cord motor neurons. Nature *355*, 838-841.

White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1986). The structure of the nervous system of Caenorhabditis elegans. Phil. Trans. R. Soc. Lond. B Biol. Sci. *314*, 1-340.

Wightman, B., Baran, R., and Garriga, G. (1997). Genes that guide growth cones along the *C. elegans* ventral nerve cord. Development *124*, 2571-2580.

Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., *et al.* (1994). 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. Nature *368*, 32-38.

Wood, W. B. (1988). The nematode Caenorhabditis elegans, W. B. Wood, ed. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press)

4.6 Addendum

Several other pieces of data were omitted during the preparation of the main body of this chapter as a manuscript submission. Below are a description of the degenerate primers used to identify other UNC-119 homologues; the sequence of the *Drosophila* DmUNC119 gene; the punctate localization of extrachromosomal *unc-119*::UNC-119::GFP transgenes; images of some specific structures that express *unc-119*::GFP; photographs of the HSN cell bodies in N2 and *unc-119* animals; and a description of two vectors designed to allow *unc-119*-directed transgene expression in the *C. elegans* nervous system.

4.7 Materials and Methods

4.7.1 mRNA analysis in Drosophila

The 5'RACE experiment in *Drosophila* was performed with the same reagents and protocols as for *C. elegans* (§2.2.6). Total *Drosophila* RNA from adults and larvae of the strain *Cy/O*, a gift from R. Hodgetts, was prepared using the same protocol described in §3.6.1, with the substitution of 100 mg of flies for the worms. The first-strand primer was MMA21, previously used for *C. briggsae* 5'RACE (§3.2.2), but which was found to be a 100% match with the corresponding *Drosophila* sequence obtained from pDP#MM121. The second strand primer was MMA29 (5'-CAGTGAATTCCACA*GTGGCTCCCACTG-3'). An endogenous *Eco*RI site is underlined, and the point at which an intron occurs in the genomic sequence is shown is indicated by an asterisk.

Reverse transcriptase PCR (RT-PCR) was performed as follows. The total *Drosophila* RNA was extracted twice using TRIzol reagent (Gibco-BRL). The first-strand primer was MMA34 (5'-GCCGGATCCATATTGCCGCCATCGTA

G-3'), and the PCR primers were MMA34 and MMA35 (5'-GCCGGATCCAAAA*ATG*GTGACTCCCGACGAGGTG-3'). The *Bam*HI sites used for cloning are underlined, and the novel start codon is italicized. The reagents from the BRL 5'-RACE kit were used, and the manufacturer's protocol followed, to a point immediately after the 70°C treatment of the first-strand reaction mixture. PCR was then performed on 3 µL of the first-strand reaction with the following conditions: (95°C, 4'; 57°C, 60"; 72°C, 2') for one cycle, and (93°C, 1'; 57°C, 45"; 72°C, 2') for 29 cycles. A weak band of the expected size (615 bp) was reamplified and cloned into *Bam*HI-digested pBluescript KS(-).

4.7.2 C. elegans UNC-119::GFP fusions

The *C. elegans* functional GFP fusion pUGF16 is described in Section 4.2. The older full-length fusion, pUGF7, was constructed by ligation of a *Hin*DIII-*Bam*HI fragment from pDP#MM016M (§4.2.1) into pTU#61 (Chalfie *et al.*, 1994). This fusion includes nearly all of the UNC-119 protein and rescues the *unc-119* phenotype.

The first *unc-119*::GFP fusion, pUGF2, was made by cloning a *Hin*DIII-*Pst*I fragment of pDP#MM016 (§2.2) into pTU#62 (Chalfie *et al.*, 1994). This fusion only encodes 101 aa of UNC-119 and does not rescue the *unc-119* phenotype.

An *unc-119*::GFP fusion containing a larger 5' flanking sequence was made as follows. A ~9 kb *Pst*I fragment of cosmid M142 was cloned into the *Pst*I site of pPD95.77 (A. Fire, J. Ahnn, G. Seydoux and S. Xu, pers. comm.), to make pUGF18. The upstream *Pst*I site is located within the M142 polylinker, and the downstream site is identical to the fusion point of pUGF12 (§4.2.1).

4.7.3 C. briggsae full-length GFP fusion

The *C. briggsae* GFP fusion pUGF8 (pDP#MM080) was made as follows. An intermediate plasmid, pDP#MM059, was constructed by subcloning an *Xba*I-*Nde*I fragment of pDP#MM035 into pUR#91 (a gift from W. Wadsworth). A *Bam*HI-*Bgl*II fragment of this plasmid was then cloned into the *Bam*HI site of pTU#61 (Chalfie *et al.*, 1994). Plasmid pUGF8 encodes 216 aa (almost all) of CbUNC-119 fused to the 'original' GFP (explained below).

4.7.4 Construction of pDP#MM065 unc-119 expression vector

The pDP#MM065 expression vector was constructed in several steps. First, a *Hin*DIII-*Ase*I fragment of pDP#MM016, containing the *unc-119* promoter and the first part of the *unc-119* cDNA, was cloned into pUR#91 (a gift from W. Wadsworth) digested with *Hin*DIII and *Nde*I, to make pDP#MM055. A slightly larger *Hin*DIII-*Bam*HI fragment was removed from pDP#MM055 and cloned into pBluescript KS(-) digested with the same enzymes, to make pDP#MM056. The now unique *Eco*RI site in pDP#MM056, present in the *unc-119* promoter, was deleted by digesting the plasmid with *Eco*RI, filling the ends with the Klenow fragment of *E. coli* DNA polymerase I (GibcoBRL), and recircularizing to produce pDP#MM057. The deletion of the *Eco*RI site was confirmed by failure to digest with *Eco*RI, and ability to digest with the novel *Ase*I site created by the filled site (5'-GAATTAATTC-3').

The *Hin*DIII-*Bam*HI insert in pDP#MM057 was removed and cloned into pUR#91 digested with *Hin*DIII-*Bgl*II, to generate pDP#MM063. This plasmid was subjected to *in vitro* mutagenesis with the oligonucleotide MMA6 (5'-CCC ATT ATC GAT AAG ATC TCC ACG GTG GCC-3') to remove a duplicate *Eag*I/*Not*I site, remove two ATG start codons, and create a *Bgl*II site and second *Cla*I site. The final vector, pDP#MM065, contains the *unc-119* promoter with a filled *Eco*RI site, upstream of 11 unique cloning sites in a pBluescript SK(+) backbone. The entire promoter and presumed transcription start site can be removed with *Cla*I.

4.8 Degenerate primers for identifying UNC-119 homologues

Degenerate oligonucleotide PCR allows the amplification of a nucleotide coding sequence from the corresponding amino acid sequence (Compton, 1990). To facilitate the identification of putative UNC-119 homologues from other phyla, degenerate oligonucleotide primers were designed based on several regions of strong conservation between UNC-119 and HRG4 (Figures 4.8 and 4.10). Of these, four were chosen for synthesis (B, C, E and F).

The intended use of the primers was to first generate PCR products using the outer primer set (B and E), and then eliminate non-UNC-119 homologues by reamplification with the nested primers (C and F), which also contain *Bam*HI restriction sites for cloning. Pilot experiments with primers B and E were able to amplify a product from genomic DNA of *C. elegans* and *C. briggsae* of expected size, and from the related nematode *C. remanei* (data not shown). However, no amplification was obtained from DNA of *H. sapiens* (where an HRG4-specific band was expected), *D. melanogaster*, the snail *Helisoma trivolvis*, and the jellyfish *Polyorchus penicillatus*. The failure to detect an HRG4-specific product in humans may have been due to prohibitive PCR product size, or the interruption of one of the primer binding sites by an intron.

Amplification was obtained with the nested primers C and F, from the DNA of the three *Caenorhabditis* species, humans, and the fruit fly *Drosophila*. Despite the degenerate nature of the primers, only a single product was recovered in each species except *C. elegans*, which generated an extra product of size ~400 bp in addition to the anticipated 694-bp *unc-119* band. The second

C. elegans product, and the human and fly bands were cloned and sequenced. The second *C. elegans* band corresponded to an artifactual product generated by the C primer alone, and contained sequences from the LG *I* cosmid T28F2. The human product overlapped the HRG4 cDNA (allowing for a single intron). The *Drosophila* product also appeared to have a single intron, and encoded a protein fragment with significant homology to both HRG4 and UNC-119 (see §4.2.4). The apparent intron in each of the *Drosophila* and HRG4 genomic PCR products is located in the same place in the *Caenorhabditis* genes (Figure 4.11).

4.9 The *Drosophila* DmUNC119 gene

The 455-bp degenerate oligonucleotide PCR product obtained from *Drosophila melanogaster* was used to isolate pDP#MM121, a 4.5-kb *Hin*DIII fragment containing a large portion of the gene (see §4.2.4). A restriction map was generated (Figure 4.11) and several smaller subclones were obtained for sequencing. The restriction map showed that pDP#MM121 contained ~2.5 kb of the fly gene. The sequence, shown in Figure 4.12, predicted a second intron which is expected to interrupt base pairing between the B degenerate primer and the genomic DNA. This accounts for the failure of the B-E primer pair to generate a product in *D. melanogaster*.

The homology between the fly coding sequences and those of HRG4 and UNC-119 allowed the deduction of the more upstream parts of DmUNC119; however, the lack of conservation upstream of region A precluded confirmation of the correct residues. As a precaution against incorrect sequence information, this region was subjected to repeated rounds of sequencing. Unfortunately, no in-frame ATG codons were found. Furthermore, an in-frame stop codon exists 186 base pairs upstream of the start of region A. Therefore, either an error persists in the sequence, or this portion of the gene is intronic (and must therefore contain a splice acceptor).

An attempt was made to identify the 5' end of the putative DmUNC119 message. 5' RACE analysis was performed as was done for the *C. elegans* and *C. briggsae* genes (§2.2.6 and §3.2.2). Two products were obtained and cloned. One of the products, pDP#MM136, started within the conserved region A (shown in Figure 4.12), and was therefore likely to be the result of a degraded mRNA or incomplete cDNA. The second product, pDP#MM135, began 173 bp from the start of region A and overlapped the genomic sequence. As there were no in-frame start codons found, either the genomic sequence is incorrect, or the clone is an artifact. The determination of the true 5' end of the putative message may have to rely on the screening of a cDNA library.

The existence of the two introns in the conserved portion of DmUNC119, predicted by alignment with HRG4 and UNC-119, was confirmed by RT-PCR (see §4.2.2). As the RT-PCR product resulted from post-transcriptional processing, this experiment demonstrated that the putative DmUNC119 locus expresses a transcript *in vivo*.

4.10 Unusual punctate localization of some *unc-119*::GFP reporters

Some *unc-119*::GFP fusions result in punctate subcellular localization of GFP. Though initially ascribed to the putative *in vivo* localization of UNC-119, this unusual localization was found to depend on the transgene constructs, rather than UNC-119 function of the fusion protein. Therefore, it was concluded that this effect was artifactual, and not necessarily related to the true *in vivo* localization of UNC-119. In this section, the characteristics of the transgenes and their expression will be discussed.

The first *unc-119*::GFP reporter fusion made during the course of this work was the pUGF2 *unc-119*::GFP fusion, which used the older GFP vector pTU#62 (Chalfie *et al.*, 1994), and contained the same upstream sequences as in the *lacZ* fusion described in Chapter 2. Vectors containing the S65C mutation in the GFP cDNA so greatly enhanced the fluorescence and stability of GFP transgenes, that they quickly superseded the original vectors (A. Fire, J. Ahnn, G. Seydoux and S. Xu, pers. comm.) The S65C GFP fusion analogous to pUGF2 is the pUGF12 plasmid, described in the main body of this chapter.

The pUGF2/pUGF12 GFP fusions retain 101 aa (of 219 aa) of UNC-119, and do not cause a dominant transgene phenotype, nor do they rescue the mutant phenotype of *unc-119* animals. Therefore, targeting of GFP fluorescence, which was observed to be diffuse in axons, was thought unlikely to reflect any localization of the endogenous *unc-119* gene product.

In order to 'tag' UNC-119, GFP fusions were constructed that retained the majority of the UNC-119 coding region. This might allow the fusion protein to rescue the *unc-119* mutant phenotype and reveal the potential subcellular localization of UNC-119. This experiment is similar in strategy to traditional "epitope tagging" approaches, in which coding regions added to a transgene allow detection of the fusion protein with commercially available antibodies (reviewed in Kolodziej and Young, 1991). However, the epitope added is typically small (less than 30 aa); with a functional GFP fusion, the coding region added to the gene under study is at least 239 aa (the size of GFP). In the GFP expression vectors designed for use in *C. elegans*, a further ~30 aa are added as a result of the vector backbone. Therefore, the total size of the 'GFP tag', approximately 270 aa, is significant and might be expected to disrupt function of a fusion transgene. In the case of full-length UNC-119, the addition of GFP more than doubles the size of the fusion protein, since UNC-119 is only 219 aa long.

A pilot experiment was performed to determine if a full-length UNC-119::GFP fusion could retain UNC-119 function. The plasmid pUGF7 was constructed, which contained the wild-type *unc-119* gene (until a region just prior to the stop codon) fused to the original GFP cDNA and the 3'UTR from the *unc-54* gene (Chalfie *et al.*, 1994). Animals homozygous for *unc-119(e2498)* were made transgenic for pUGF7 and found to be rescued for locomotion and dauer formation, consistent with function of the UNC-119::GFP fusion. Under fluorescence, GFP signal was still present (consistent with formation of a functional chromophore by the UNC-119::GFP fusion). Therefore, the UNC-119::GFP fusion protein retained properties of both UNC-119 and GFP (assuming that the fusion protein remains intact *in vivo*).

While GFP signal was still observed in axons, however, it took on a punctate appearance (Figure 4.13A). These punctate regions were not nuclei, as demonstrated by doubly staining the cells with DAPI to show DNA. Furthermore, immunological staining with an anti-synaptotagmin antibody showed that the punctate regions were not synaptic vesicles (C. Link, pers. comm.; Nonet *et al.*, 1993).

In addition to the unusual punctate expression within axons, there was a pattern of GFP expression consistent with targeting of GFP to muscle dense bodies in the head (Figure 4.13B). Simultaneous detection of dense bodies and UNC-119::GFP was not performed, so it is not clear whether the fluorescence was truly within the dense bodies, or merely associated with them. The localization of GFP to a muscle tissue was unexpected, as there are no apparent muscle defects in *unc-119* mutants. Muscle expression could have been the observed non-neuronal staining seen in the *lacZ* fusions described in Chapter 2.

At this point, it seemed as if the true *in vivo* localization of UNC-119 had been determined, as there was a correlation between the rescue of the *unc-119* defects and the punctate staining. However, as compelling as these preliminary results were, several additional lines of evidence cast doubt on their validity.

First, two *C. briggsae* UNC-119::GFP fusions gave punctate expression in *C. elegans* (pUGF15 and pUGF8; §3.8). The two constructs differ in the portions of UNC-119 retained in the fusions: pUGF15 encodes 130 aa of CbUNC-119, and pUGF8 encodes 216 aa of CbUNC-119, while the original *C. elegans* pUGF2/pUGF12 fusions are predicted to encode only 101 aa of UNC-119 (Figure 4.14). The difference in expression between pUGF12 and pUGF15/pUGF8 seemed to suggest that the additional amino acids (i.e. positions 101-130) were responsible for the localization.

Second, a full-length HRG4::GFP fusion targeted GFP to axons in a diffuse manner (§4.3.9). As the HRG4::GFP transgene was also capable of rescuing the *unc-119* defects, punctate localization of this UNC-119 homologue is therefore not correlated with function.

Third, a functional UNC-119 fusion to the S65C GFP variant, similar to pUGF7, appeared to lose the punctate expression. When this plasmid (pUGF16) was chromosomally integrated, all five of the independent lines obtained showed diffuse axonal staining of GFP and no dense body fluorescence (Figure 4.7). The differences between pUGF7 and pUGF16 were the S65C mutation in GFP, the presence of novel synthetic introns in the GFP coding region, and the use of the *unc-119* 3'UTR instead of the 3'UTR from *unc-54* (§4.2.1; A. Fire, J. Ahnn, G. Seydoux and S. Xu, pers. comm.).

The simplest explanation for the punctate localization of some UNC-119::GFP transgenes is that it is an artifact that depends on the transgene. This may be because some transgene products might be more stable, and therefore accumulate the protein to higher levels. There is precedent for GFP-dependent targeting in prokaryotes: the mutated forms of GFP were found to be more soluble in bacterial cytoplasm, while the original GFP accumulated in non-fluorescent inclusion bodies (Cormack *et al.*, 1996). By extension, the punctate fluorescence may simply be the result of transgene overexpression combined with high protein stability.

As a final note, all *unc-119* mutant strains carrying a functional UNC-119::GFP reporter were observed to have a wild-type phenotype, suggesting that the punctate localization of GFP does not have drastic consequences for nervous system development.

4.11 Structures visualized with *unc-119*::GFP

The high resolution enabled by GFP fusions allows the identification of particular structures expressing a particular transgene (Chalfie *et al.*, 1994). During the study of *unc-119*::GFP expressing strains, some specific structures allowed identification by morphology and location; images of these have been grouped in this section. Three cells identified to express *unc-119*::GFP included the mechanosensory PVM and ALM neurons, and the canal-associated neuron (CAN) as shown in Figure 4.15. These neurons are particularly easy to identify because of the lateral placement of their cell bodies (White *et al.*, 1986). It is worth mentioning that when these cells were observed in an *unc-119* mutant background, they were always observed in their wild-type location. This suggests that despite the apparent expression of *unc-119* in these cells (implying a requirement for UNC-119), there were no obvious cell migration or axonal defects (data not shown).

The male tail neurons, whose nuclei were postulated to accumulate Xgal stain from an *unc-119*::lacZ fusion (Chapter 1), can also be visualized with *unc-119*::GFP (Figure 4.16A). A putative male mating defect, anticipated by expression of *unc-119* in these neurons, could not be assessed because of the strong locomotory Unc defect.

An unanticipated cell type was identified by its ability to accumulate *unc-119*::GFP. Fixation of *unc-119*(*e2498*) animals carrying the *edIs6 unc-119*::GFP integrant allowed individual gonad arms to be seen (the result of cuticle disruption). In 3/3 gonad arms visualized, the distal tip cell had accumulated GFP (Figure 4.16C). The implications of this expression will be discussed in Chapter 6.

Lastly, accumulation of high amounts of GFP was observed in coelomocytes (Figure 4.16D). There are six coelomocytes in the hermaphrodite and five in the male, situated in the coelomic cavity (reviewed in White, 1988). These cells are not necessary for survival, and likely have a scavenging role (White, 1988). Therefore, either *unc-119*::GFP is expressed in coelomocytes, or it accumulates in these cells from the coelomic cavity.

4.12 Visualization of HSN cell bodies in N2 and *unc-119*

The failure of *unc-119* animals to respond to imipramine was consistent with a presynaptic defect (see §4.3.2). The HSN cell bodies were visualized in *unc-119* animals using an anti-serotonin antibody and the procedure described in Desai *et al.* (1988) and §4.2.7. In Figure 4.17, one of each pair of the HSNs is shown; *unc-119* mutants (5/5 animals) showed a normal mid-body location for both HSN cell bodies compared with wild type. Therefore, a migration defect is ruled out as a basis for the egg-laying phenotype (Desai *et al.*, 1988).

4.13 Vectors for transgene expression from the *unc-119* promoter

The apparent pan-neural expression of *unc-119* reporters, and the early timepoint in development at which expression begins, suggested that the *unc-119* promoter would be useful to other researchers as a means by which any coding region could be expressed in the *C. elegans* nervous system. To make the promoter versatile to such applications, the *unc-119* promoter was modified from the original pDP#MM016 plasmid containing the entire *unc-119* gene (Chapter 2). An *Eco*RI site was removed, and eleven unique cloning sites were added downstream of the transcription start site to create pDP#MM065, an *unc-119* expression vector in a pBluescript SK(+) backbone (Figure 4.18A).

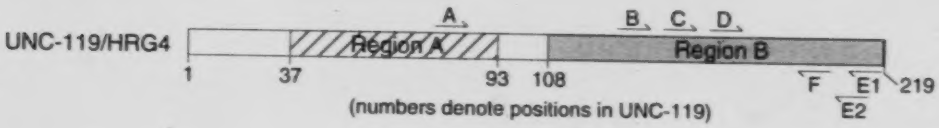
The function of the promoter in pDP#MM065 was tested by removing it with *Hin*DIII and *Bgl*II and cloning into the GFP vector pPD95.75 digested with *Hin*DIII and *Bam*HI (A. Fire, J. Ahnn, G. Seydoux and S. Xu, pers. comm.). Animals transgenic for this plasmid, pUGF9, showed neural expression similar to pUGF12, although it was reduced in intensity (data not shown).

The pDP#MM065 vector, therefore, allows cloning of complete cDNAs that have a 3'UTR (as none is present in pDP#MM065), or of promoterless genomic sequences. A 3'UTR has since been added (R. Baumeister, pers. comm.): A *Hin*DIII-*Bam*HI fragment of pDP#MM065 was introduced into the same sites in pPD49.26, a general expression vector with a pUC19 backbone, which was the same vector used to create the GFP and *lacZ* fusion vectors described earlier (A. Fire, pers. comm.). The resultant plasmid, pBY103, includes the *unc-119* promoter and start of transcription, many unique restriction sites, a synthetic intron, and the 3'UTR of the *unc-54* gene (Figure 4.18B). As such, this vector is suitable for cloning in cDNA fragments or incomplete genomic DNAs.

As of the writing of this thesis, the *unc-119* promoter has already been useful in the study of nervous system genes in *C. elegans*. This includes the *unc-119*-directed expression of β-peptide (C. Link, pers. comm.), UNC-6 (W. Wadsworth, pers. comm.), and VAB-1 (A. Chisolm, pers. comm.).

4.14 An *unc-119*::GFP fusion with a larger 5' flanking region

All of the *C. elegans unc-119* transgene fusions described contain approximately 1.2 kb of 5' flanking sequence. While only limited similarity was found to exist between the *C. elegans* and *C. briggsae* promoters, very close to the transcription start sites (Chapter 3), it is possible that there are other *cis*-acting regulatory elements farther upstream. However, no obvious changes in expression pattern were seen with a GFP fusion plasmid, pUGF18, constructed using approximately 7 kb of upstream sequence. This fusion produced similar neural expression as the other *unc-119* reporters, although the GFP fluorescence was reduced in intensity (data not shown).



A consensus VLFEI$\overset{A}{\underset{K}{}}$KP

	V	L	F	E	I	A K	K	P	
5'-	GUU	UUA	UUU	GAA	AUU	GCU	AAA	CC	- 3'
	C	G	C	G	C	C	G		
	A	CUU			A	A			
	G	C				G			
		A				AAA			
		G				G			

sequence: 5'- GTI YTI TTY GAR ATH RMI AAR CC -3'

B consensus VGATVEF

(MMA22)

	V	G	A	T	V	E	F	
5'-	GUU	GGU	GCU	ACU	GUU	GAA	UUU	- 3'
	C	C	C	C	C	G	C	
	A	A	A	A	A			
	G	G	G	G	G			

sequence: 5'- GTI GGI GCI ACI GTI GAR TTY -3'

C consensus FRMIERH

(MMA23)

	F	R	M	I	E	R	H	
5'-	UUU	AGA	AUG	AUU	GAA	AGA	CAU	- 3'
	C	G		C	G	G	C	
		CGU		A		CGU		
		C				C		
		A				A		
		G				G		

sequence: 5'- ga cgg atc cGI ATG ATH GAI MGI CAY -3'
BamHI

Figure 4.10 Design of degenerate PCR primers to amplify putative homologues of UNC-119/HRG4 (continued on next page). The top figure shows the locations of the various primers within the UNC-119 coding region. Only the B, C, E2 and F primers were synthesized. The *Bam*HI sites used for cloning are indicated. Nucleotide abbreviations: I = inosine; R = A or G; Y = C or T; S = G or C; W = A or T; H = A or C or T; N = A or C or G or T.

D consensus FDF$^{E}_{H}$FGFC

```
                 E
     F   D   F   H   F   G   F   C
5'- UUU GAU UUU GAA UUU GGU UUU UG - 3'
      C   C   C   G   C   C   C
                CAU       A
                  C       G
```

sequence: 5'- TTY GAY TTY SAI UUY GGI UUY UG -3'

E1 consensus KADYSY

```
     K   A   D   Y   S   Y
5'- AAA GCU GAU UAU AGU UAU - 3'
      G   C   C   C   C   C
          A         UCU
          G           C
                      A
                      G
```

complementary sequence: 5'- RTA ISW RTA RTC IGC YTT -3'

E2 consensus VMHNKAD
(MMA24)

```
     V   M   H   N   K   A   D
5'- GUU AUG CAU AAU AAA GCU GA - 3'
      C       C   C   G   C
      A                   A
      G                   G
```

complementary sequence: 5'- TC IGC YTT RTT RTG CAT NAC -3'

F consensus DSFYFV
(MMA25)

```
     D   S   F   Y   F   V
5'- GAU AGU UUU UAU UUU GU - 3'
      C   C   C   C   C
        UCU
          C
          A
          G
```

complementary sequence: 5'- ct ggg atc cAC RAA RTA RAA ISW RTC -3'
BamHI

Figure 4.10, continued.

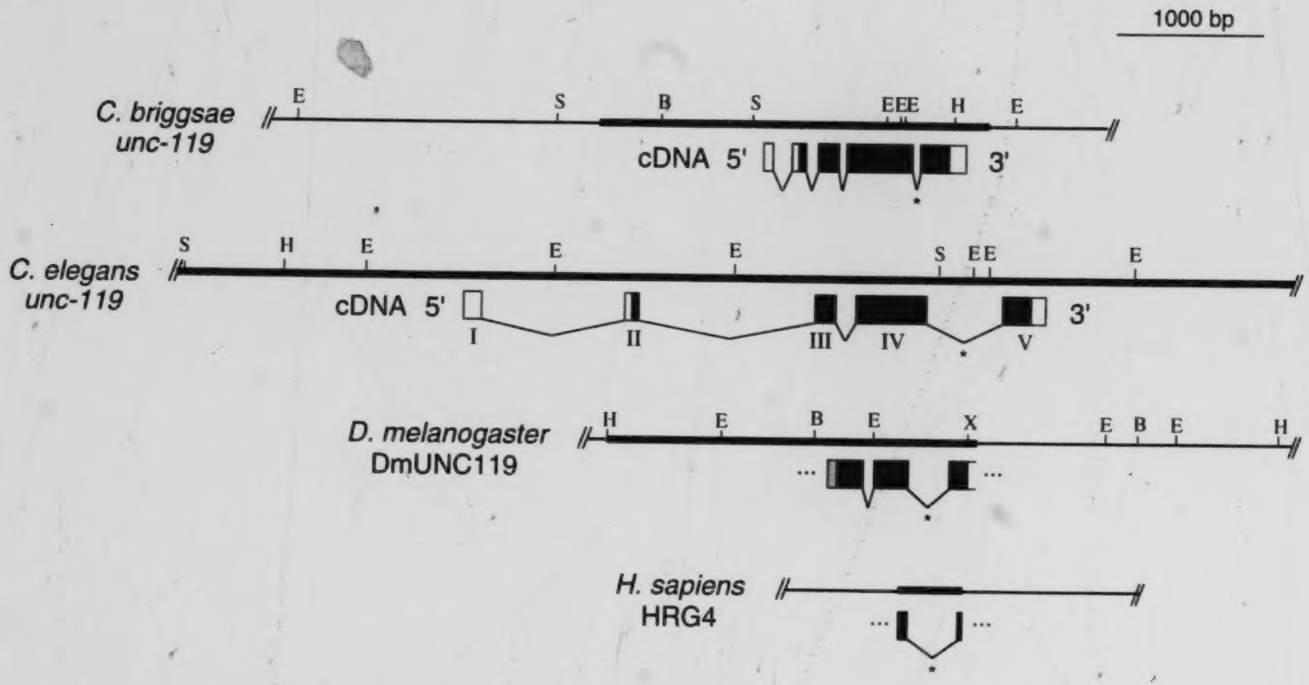


Figure 4.11 Gene arrangements of *C. elegans* and *C. briggsae unc-119*, *Drosophila* DmUNC119, and human HRG4. Thicker regions in the genomic DNA indicate segments that have been sequenced. Black segments in the cDNAs indicate the coding regions. In the *D. melanogaster* cDNA, a lighter shaded portion of the cDNA indicates upstream codons in frame with conserved amino acids further downstream. Unknown cDNA ends are indicated with an ellipsis (...). The introns indicated by an asterisk (*) occur in the same relative position across all of the homologues. B, *Bam*HI; E, *Eco*RI; S, *Sst*I; X, *Xba*I.

```
aagctttctgcaaacgttcgcttactttcggttatacaataaataatgttaaaatctgaatcgcctttttctagatgtaaaatgtagaatattttgaa 100
tattgcgccaacagtgtgacccctgctagaacgcagaaatgtttttttgcgagtatgtggcaacgttggtcgtaagtttttccccatgcgcgaacggtct 200
caccgatttcctgtaatgcatcatcgataatctatcgattgctgagtag↓tgcaaaaaaagtgggcaaaaaaggtgcaatgtttagcgtttaagtgcgtta 300
agttgcttttgcggacgacaacaacaacaccagcagcattgccggtcaaattgatctaattgatttcgggcagtgcaacagcaactaaaaccaaattaa 400
taaatgaatccctgcgcatacataaacatataaaaacacacacacacacatacattcgcataaaggacacgaaccgaaaagggaaaaacaaacacaaa 500
ggaaaaaaaaaggaaaaaggagagcgcagcttttttcagaaattgggtcatttgttgttgttgttgttgttgctgctgttgctggtgatttcaaacg 600
aatttcatttgttgctgtcgtcgtcggttcggtgttgttgttgttgttgctgttgttgcaatggtcgtggacgttgttgtcgctgacagctaaatgttat 700
ggcaacacttgcaacctgaattcgtgtgggtaagtagtgtttgcatttatgttgaatttattttaacttatataccttatacttaacttattttttgtgct 800
tttcgccaaataactattcatatttggcaaaaattaatagtgaggagcggaaagcaaaggataatagggtagcggagacgaagccttatatttccaccag 900
aaaatcctggcacaatttatgacaactctacacttgtgtgtgtttgtgttggtgagtaacttaacaactatgacttctttttttttttatgttaaaaaa 1000
attgcatatttcattattcagttagttgagtatagaatcaaagtattgtgctatattatattactcctataaccaatgctattcctttctattcgcagag 1100
tgaagacagcggagtagcttcttcaaatcaaaggataccactgcagcagccatccagccaggagcttaaaaaagctgttcagcttgagaaatccgaaacg 1200
                                      in-frame stop codon ?           ↓ 5'RACE product start (pDP#MM135)
cagactatagacaaataataccaccagattcattccgtgcacaatgagcgtggtgggtaagcaactaaatccggtgcaatctccggtgcaggcgcagtga 1300
caacgtcttcctccgcggccgcaggatcctcttcttcgaacagtggagtggaggccaacggcggcagtggaggatcgtcgggagctgcagcagcgggcgc 1400
                                        ... V  T  P  D  E  V  L  H  L  T  K  I  T  D  D  Y  L  C  S
tggtgcctccggtgatgccaaacggcctgcggagtcttccagcGTGACTCCCGACGAGGTGCTCCATCTGACCAAGATCACGGACGACTATCTCTGCTCC 1500
 A  N  A  N  V ↓F  E  I  D  F  T  R  F  K  I  R  D  L  E  S  G  A  V  L  F  E  I  A  K  P  P  S  E  Q
GCCAATGCAAATGTGTTCGAGATTGATTTCACGCGGTTCAAGATTCGCGACCTGGAGAGCGGCGCTGTGCTCTTCGAGATCGCCAAACCGCCGAGTGAAC 1600
   Y  P  E  G  L  S  S  D  E  T  M  L  A  A  A  E  K  L  S  L  D  D  T  A  D  P  N  A  G  R  Y  V  R
AATATCCGGAAGGACTGTCCTCGGATGAAACCATGCTGGCGGCTGCCGAGAAATTGTCACTGGATGACACTGCCGATCCAAATGCCGGACGCTATGTGCG 1700
```

Figure 4.12 Incomplete nucleotide sequence and predicted amino acid sequence of the DmUNC119 locus, obtained from the *Drosophila* clone pDP#MM121. Predicted cDNA sequences are shown in capital letters, and the predicted amino acid sequence is denoted by single-letter abbreviations above the cDNA sequence. The 5' ends of the two 5' RACE products are indicated with arrows. An apparent in-frame stop codon, upstream of the conserved sequences, is shown with a question mark. (Continued on next page)

```
  Y  Q  F  T  P  A  F  L  N  L  K  T  V  G  A  T
CTATCAGTTCACACCGGCATTTCTCAACCTCAAAACAGTGGGAGCCACgtaggttatacagtatttgcatttgttttaaatgcctctcaatttgtaagtt 1800
                V  E  F  T  V  G  S  Q  P  L  N  N  F  R  M  I  E  R  H  F  F  R  D  R  L  L  K  T  F
tcccttatttgcagTGTGGAATTCACTGTGGCAGCCAGCCGCTTAATAATTTTCGTATGATCGAACGCCACTTCTTCCGCGATCGCCTGCTAAAGACGT 1900
   D  F  E  F  G  F  C  F  P  F  S  K  N  T  V  E  H  I  Y  E  F  P  N  L  P  P  D  L  V
TTGACTTTGAGTTCGGCTTTTGCTTTCCATTTTCGAAGAATACGGTTGAGCACATCTACGAGTTTCCTAACCTTCCACCCGACCTAGgtgagcaattttt 2000
acataatcccagccagcatgacacgcaattgaaagttgctgccaaaagttgaaggggtttgtggggtcaaggtcgtttacgatttgcgcctgtgagatga 2100
ggctctattggtcagattgcgagcagctatcatggttttatttttgttatcggtttgtaggatgtggacttattaatgaaatgcaattcacataaata 2200
                               A  E  M  I  S  S  P  F  E  T  R  S  D  S  F  Y  F  V  G  N  R  L  V  M
tgttcgttcgttttccatttttttttagTTGCTGAGATGATATCGAGCCCCTTTGAGACGCGCTCGGACAGCTTTTACTTTGTGGGAAACCGACTCGTCA 2300
   H  N  K  A  D  Y  A  Y  D  G  G  N  I  V
TGCACAACAAAGCCGACTACGCCTACGATGGCGGCAATATAGTCTAGAAACTCAACTATAATGGAACTGCAACTAGAGCAAAACTAATAACTAAAGAACT 2400
AGAAACTACAGAAGCTGTGAGACAAAATCAACGACTTGTCACACACACACACACACGAAACACACACACGTAAA 2474
```

Figure 4.12, continued. Sequences which cross-hybridize to the degenerate primers C and F are boxed, which indicate the ends of the cross-species PCR product obtained. The 3' end of the cDNA has not determined, as the known sequence does not include a polyadenlyation site. The stop codon is underlined.

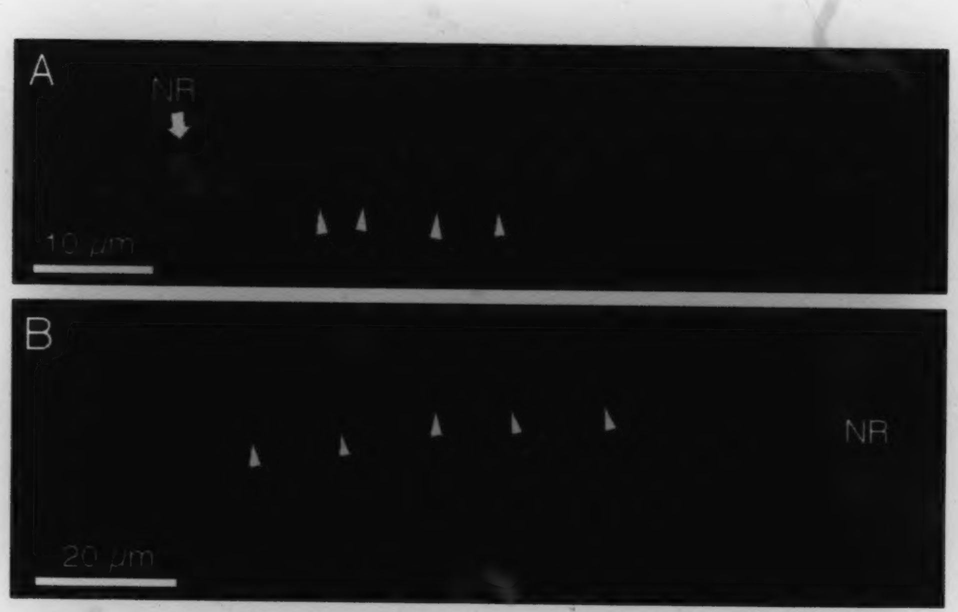


Figure 4.13 *unc-119(e2498)* animals transgenic for the functional UNC-119::GFP fusion pUGF7. Animals were fully rescued for the *unc-119* locomotory defect. **A**, Punctate fluorescence is observed along the entire length of the animal in the nervous system, including the ventral nerve cord (arrowheads) and the nerve ring (NR). (Compare to the diffuse axon fluorescence in Figure 4.5.) **B**, Punctate staining in the head region that resembles the location of muscle dense bodies (compare Figure 3.10). Fluorescence in the nerve ring (NR) is visible below the plane of focus. Anterior is to the left in both panels.

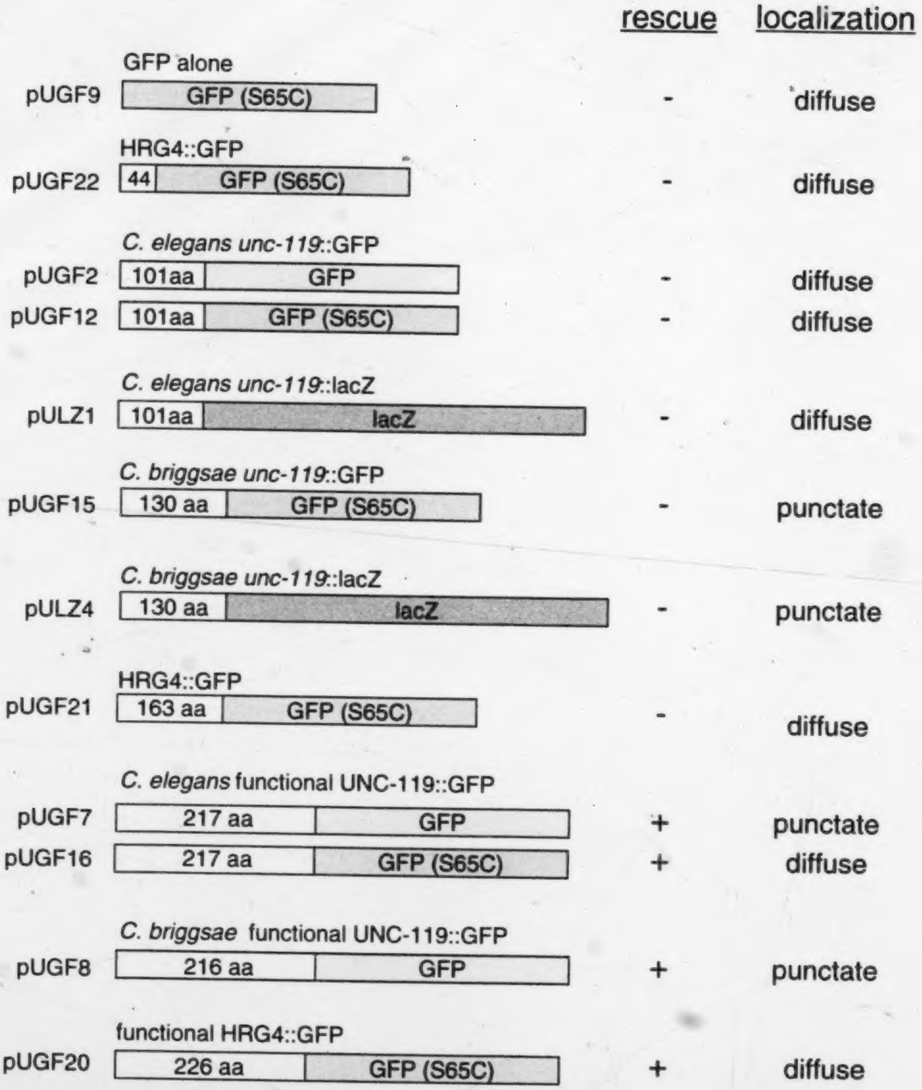


Figure 4.14 Comparison of localization of various *unc-119*::reporter fusions. Rescue denotes the ability of the fusion protein to rescue the *unc-119* locomotory defect. 'Punctate' expression refers to both punctate fluorescence in axons and apparent dense body fluorescence in the head; 'diffuse' refers to uniform axon fluorescence. There is no strong correlation between a particular fusion and the type of localization. All fusion proteins were driven by the same *unc-119* promoter fragment, except for the *C. briggsae* fusions, which used the *C. briggsae* promoter. Fusion plasmids that rescue the *unc-119* phenotype were injected alone; others were coinjected with either pRF4 or an *unc-119* rescuing plasmid (pDP#MM016B). The depicted sizes of the fusion proteins are approximately to scale, except for lacZ, which is much larger than shown.

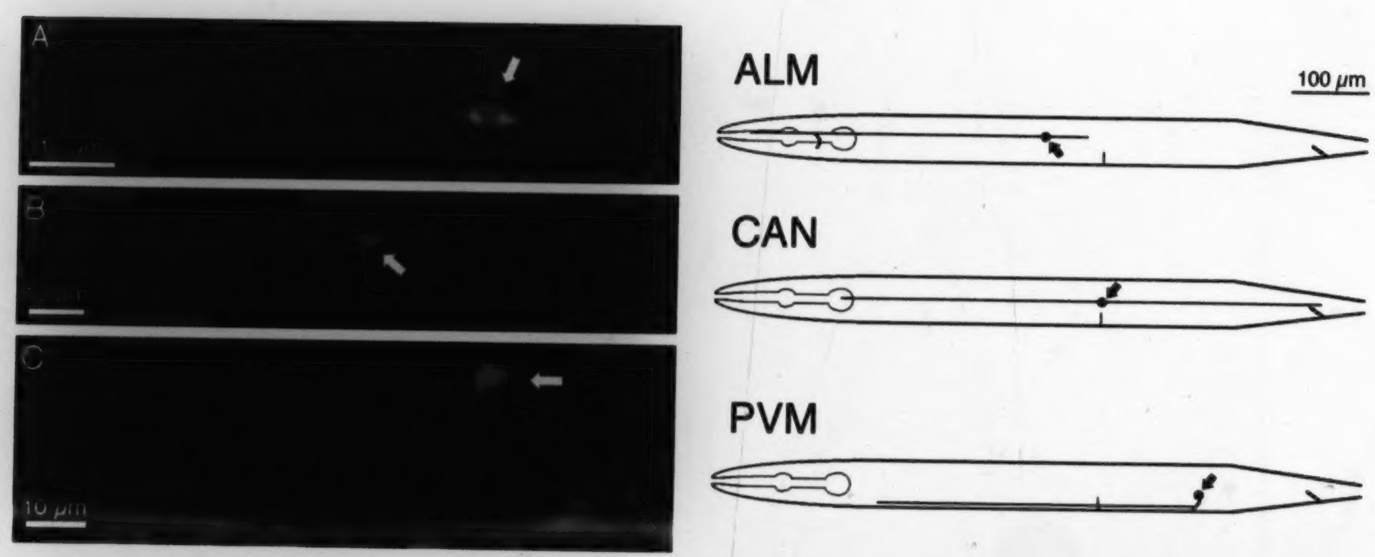


Figure 4.15 Some specific neurons expressing *unc-119*::GFP from the pUGF12 fusion. Neurons were identified based on morphology and location. Line drawings of each neuron type, shown on the right of each image, were redrawn from White *et al.* (1988). Cell bodies are indicated by arrows. **A**, *unc-119*(*e2498*) animal transgenic for pUGF12 and pRF4 showing expression in ALM neuron. **B**, *unc-119*(*ed3*) animal transgenic for pDP#MM103 (*unc-119*::HRG4 fusion) and pUGF12 showing fluorescence in CAN neuron. **C**, *unc-119*(*ed9*) animal transgenic for pDP#MM016B (*unc-119* rescuing clone) and pUGF12 showing expression in PVM neuron.

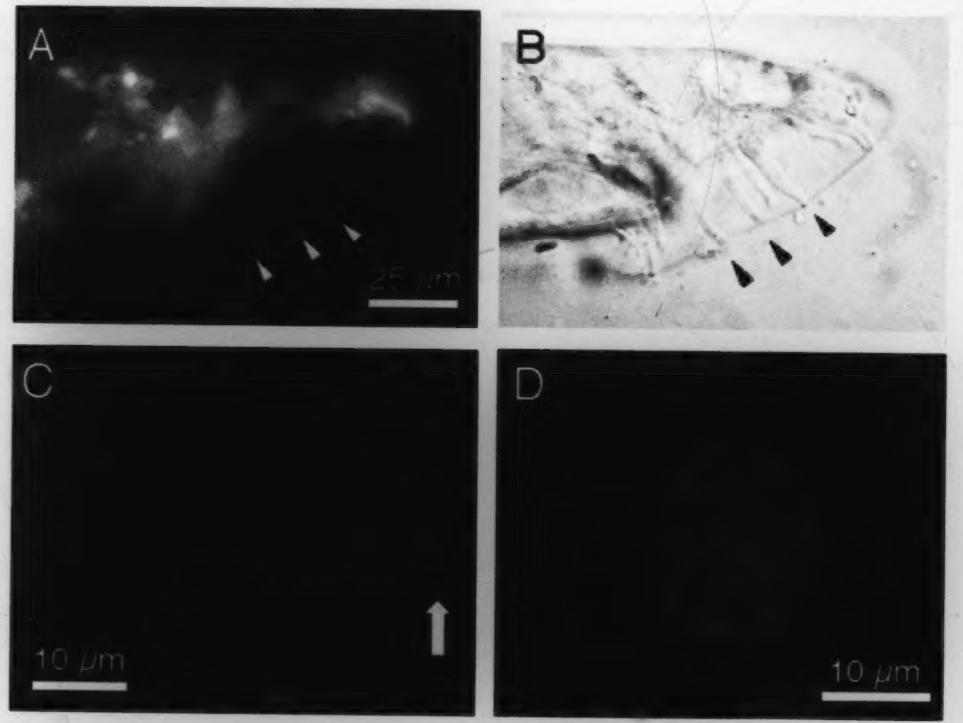


Figure 4.16 Accumulation of *unc-119*::GFP (from pUGF12) in particular cells. **A**, *unc-119*(*ed9*) animal carrying pUGF12 and pDP#MM016B (*unc-119* rescuing plasmid) showing accumulation of neurons in the male tail (arrowheads). **B**, Nomarski image of the tail in (A). **C**, GFP accumulation in the distal tip cell (DTC) of an isolated gonad of an *unc-119*(*e2498*) animal carrying *edIs6* (integrant of pUGF12). Although the image has been electronically enhanced, it may be difficult to differentiate the GFP (arrow) from the autofluorescence of the gonad. **D**, Accumulation of GFP in a coelomocyte of the same strain as the animal in (A).

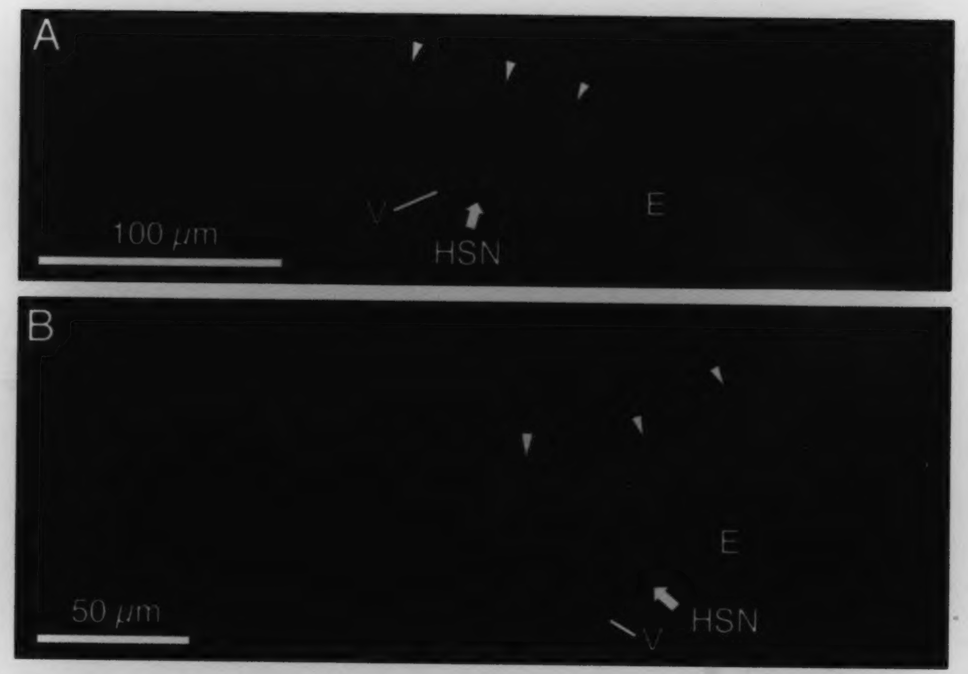


Figure 4.17 Staining of HSNs with an anti-serotonin antibody. Although this procedure has been reported to stain the entire HSN process (Desai *et al.*, 1988), only the cell bodies were visible. Only one of the two cell bodies is shown here. **A**, Wild type. **B**, *unc-119(e2498)*. In both strains, the HSN cell bodies were located just posterior to the vulva (V). The distal part of one gonad arm in the dorsal half of each animal(arrowheads) and an embryo located ventrally (E) are indicated. Anterior is to the left in both panels.

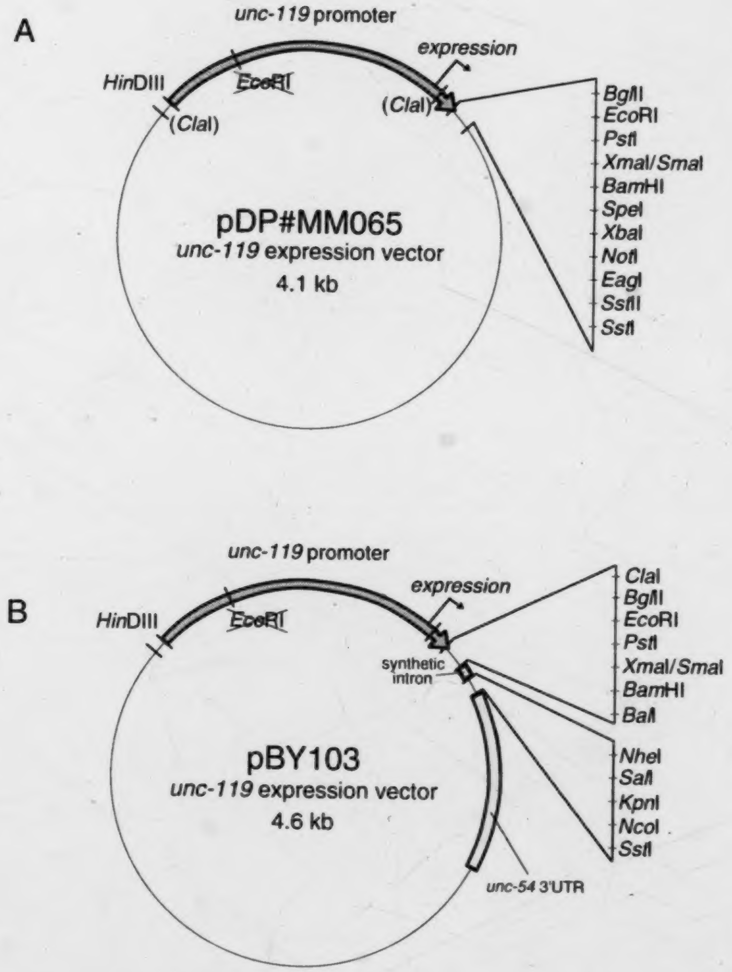


Figure 4.18 *unc-119* expression vectors which allow transgene expression in the nervous system. All restriction sites shown are unique, except for the *Cla*I site in pDP#MM065. **A**, pDP#MM065 vector, suitable for cloning of complete cDNAs or promoterless genomic fragments, as there is no 3'UTR present. This vector is in a pBluescript SK+ backbone. **B**, pBY103 vector, in a pUC19 backbone (cloned by R. Baumeister, Munich). A synthetic intron and universal 3'UTR (from the *unc-54* gene) allow the cloning of incomplete cDNAs or genomic fragments. Both vectors encode ampicillin resistance.

4.15 Bibliography

Chalfie, M,. Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. Science *263*, 802-805.

Compton, T. (1990). Degenerate Primers for DNA Amplification. in PCR Protocols: A Guide to Methods and Applications. (eds. Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J.) Academic Press, Inc., San Diego, CA., pp.39-45.

Cormack, B., Valdivia, R. H., and Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). Gene *173*, 33-38.

Desai, C., Garriga, G., McIntire, S. L., and Horvitz, H. R. (1988). A genetic pathway for the development of the *Caenorhabditis elegans* HSN motor neurons. Nature *336*, 638-646.

Kolodziej, P.A. and Young, R.A. (1991). Epitope Tagging and Protein Surveillance. Methods Enzymol. *194*, 508-519.

Nonet, M. L., Grundahl, K., Meyer, B. J., and Rand, J. B. (1993). Synaptic function is impaired but not eliminated in *C. elegans* mutants lacking synaptotagmin. Cell *73*, 1291-1305.

White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1986). The structure of the nervous system of the nematode *C. elegans*. Phil. Trans. R. Soc. Lond. B Biol. Sci. *314*, 1-340.

5 Preliminary studies of the UNC-119 protein

5.1 Introduction

5.1.1 Previous work

A complete understanding of the function of the *unc-119* gene must ultimately include a description of the properties of the UNC-119 protein itself. To a certain extent, many of these may be described in terms of the *unc-119* messenger RNA, whose expression is a prerequisite for the production of the protein. A rapid approach toward understanding the spatial and temporal properties of *unc-119* expression, therefore, has been the examination of reporter genes. In particular, fusions of *unc-119* to either *lacZ* or GFP show expression in the nervous system (Chapters 2,3,4). This expression pattern is consistent with the apparent widespread neuronal defects seen in *unc-119* mutants, which is taken as strong support for the hypothesis that the native *unc-119* mRNA accumulates in neurons. Furthermore, the localization of functional UNC-119::GFP and HRG4::GFP fusions suggest that the native UNC-119 is found in neural cytoplasm, implicating a role in axons (Chapter 4).

The sequence of a protein may also provide clues as to its function. An open reading frame (ORF) can be compared with other known proteins in electronic databases by computer algorithm (Altschul *et al.*, 1997). Similarities with other known proteins may then allow deduction of enzymatic, localization or interaction properties. For example, the sequence of UNC-51 was found to be similar to serine/threonine kinases, which along with the mutant phenotype of *unc-51* animals, anticipates a role for the gene product in transducing signals required for axon outgrowth (Ogura *et al.*, 1994).

Unfortunately, neither UNC-119 nor its homologues showed significant similarities to proteins of known biochemical function, nor were any localization or structural motifs apparent (Chapters 2,4). Nonetheless, several predictions about its characteristics were made based on molecular genetic data. Ultimately, these can only provide a framework for biochemical verification of the true *in vivo* behaviour of UNC-119. To this end, several projects were initiated to create a set of resources for future experiments. This chapter describes the construction of these tools and summarizes the preliminary data obtained, with the long-term goal of answering three basic questions about UNC-119, described below.

5.1.2 Where is the UNC-119 protein found in the cell?

The determination of the endogenous *in vivo* localization of UNC-119 will require an antibody to the protein. One common method for obtaining an antibody is to obtain large amounts of a recombinant version of the protein expressed in *E. coli*, and to inject these into a suitable host animal such as rabbits or mice (Harlow and Lane, 1988). I will discuss the construction and expression of *E. coli* plasmids that encode recombinant fusions of UNC-119, the *C. briggsae* homologue CbUNC-119, and the human homologue HRG4. An attempt to generate antibodies to synthetic peptides generated from HRG4 sequence will also be described.

An alternative to generating an antibody is to covalently add an epitope that can be detected by commercially available antisera (reviewed in Kolodziej and Young, 1991). The methodology is similar to the construction of a functional GFP fusion (Chapter 4), which involves the addition of a coding sequence for the epitope into a clone of the wild type gene. I will describe the construction of plasmids encoding hemagglutinin (HA) tagged versions of UNC-119, and the results of preliminary staining attempts, which support the localization of UNC-119 to axons.

5.1.3 When is UNC-119 function required?

The knowledge of the true localization of UNC-119 will not directly answer questions about its temporal requirements in the animal. Proteins can theoretically be made long before they are required, or persist long after they are needed; function may require a second factor which may affect activity. A genetic approach to assessing the temporal requirements for a particular gene product exploits the existence of conditional alleles: the activity of a gene depends upon an external, controllable factor, the most common of which is temperature. For example, two temperature-sensitive alleles are known to exist for the *C. elegans fem-2* gene, which have been used to define a temperature-sensitive interval that presumably defines absolute requirement of FEM-2 function (Kimble *et al.*, 1984). Unfortunately, no conditional alleles of *unc-119* exist.

A more generally applicable approach is the use of a tissue-nonspecific promoter that can be controlled by temperature. The *hsp16-2* promoter from a *C. elegans* heat-shock response locus allows this type of control (Stringham *et al.*, 1992). The construction of *unc-119* heat-shock plasmids will be described, and heat-shock driven expression of a tripartite *hsp16-2*::UNC-119::GFP fusion transgene will be demonstrated.

5.1.4 With what other proteins does UNC-119 interact?

An understanding of the intracellular role of UNC-119 would include a description of interactions mediating UNC-119 function. A classic genetic approach involves a mutagenic suppressor screen. Such approaches have proved fruitful in *C. elegans*. For example, rare suppressors of the *unc-22(s12)* mutation map to *unc-54*, suggesting these proteins interact directly (Moerman *et al.*, 1982). However, as there are no known hypomorphic alleles of *unc-119* that might allow screening for extragenic suppressors, another nongenetic approach is required.

Currently, the most powerful method that allows identification of candidate interacting proteins is the yeast two-hybrid system (Fields and Song, 1989; Chien *et al.*, 1991). This system has a similar advantage to a genetic suppressor screen, in that the approach itself is open-ended, requiring no *a priori* assumptions about interactions. For example, a two-hybrid screen was used to demonstrate interaction between UNC-51 and UNC-14 (Ogura *et al.*, 1997). This approach will likely be the most rapid method of identifying UNC-119 interactions. The construction of a yeast UNC-119 bait plasmid, and results of a preliminary screen, will be described.

5.2 Materials and Methods

5.2.1 Plasmid nomenclature

As there were often many plasmids made for a particular set of experiments, I will use three-letter acronyms to faciliate reference to clones of a particular set. None of these were given corresponding pDP#MMn assignments; consequently, to allow laboratory assignment of these plasmids as per standard *C. elegans* nomenclature (Horvitz *et al.*, 1979), all clones referred to in this chapter should be prefixed as "pDP#MM," but are not written as such to avoid extra characters. (For example, pDP#MMUFP2 is abbreviated pUFP2.) The following acronyms were used to identify various recombinant plasmids: **pUFP**, unc-119 fusion protein; **pUHS**, unc-119 heat-shock; **pUHA**, unc-119 hemagglutinin tag; **pUBP**, unc-119 bait plasmid; **pUGF**, unc-119 GFP fusion.

5.2.2 Construction of GST fusion plasmids

An overview of the GST fusion plasmid inserts is shown in Figure 5.1. Plasmid pUFP1 was constructed as follows. The full-length *unc-119* cDNA clone cDNA3 was digested with *Pst*I, and the larger fragment was recircularized. A 580-bp *Bam*HI-*Sal*I fragment was cloned into similarly digested pGEX1λT.Sal. This vector is a modified form of pGEX1λT (Pharmacia) into which an *Eco*RI-*Sal*I-*Eco*RI linker has been inserted (a gift from P. Jäckle-Baldwin).

Fusions pUFP2-pUFP4 made use of a modified *unc-119* cDNA (cDNA4), in which the full-length *unc-119* cDNA was mutagenized *in vitro* to create a novel *Bam*HI site immediately upstream of the initator ATG (primer MMA7, 5'-GCT ACA ACA GGA TCC ATG AAG GC-3'). A *Bam*HI-*Eco*RI fragment encoding all of UNC-119 was cloned into the same sites in pGEX1λT to create pUFP2. The insert in plasmid pUFP3 was obtained from *Bam*HI-*Eco*RI digestion of a PCR product obtained from cDNA4 using the primers MMA5 (5'-GGA GCA TAG GAA TTC TTG AGT GAT TCC-3') and MMA7. pUFP4 was made by cloning a ~350 bp *Bam*HI-*Cla*I fragment of cDNA4 into the vector backbone of pUFP1 digested with the same enzymes. The sequence of the lefthand cloning junction and ~200 bp of each insert of pUFP1-pUFP4 was confirmed using Sequenase 2.0 with the primer 5'-GCA TGG CCT TTG CAG GG-3'.

The fusion pUFP5 (to HRG4) was made by cloning the insert of pDP#MM092 (see §4.2.2) into the *Bam*HI site of pGEX1λT. Plasmid pUFP6 was made by liberating the vector backbone of pUFP5 with *Bam*HI, and cloning in a *Bam*HI fragment from cDNA5B, an *in vitro* mutagenized derivative of cDNA4 containing a novel *Bam*HI site at the 3' end of the UNC-119 ORF (primer MMA4, 5'-CTA CTC GCA GGA TCC A*TA A*TT TCC CG-3').

For the *C. briggsae* fusion pUFP7, an RT-PCR product was obtained from total *C. briggsae* RNA using the primers MMA26 (5'-ATC ACA ACG GGA TCC GAG CTT G-3') and MMA27 (5'-TAT TTA TGG ATC CTA TGA GTA GTC-3'). The product was digested with *Bam*HI and cloned into the same site in pGEX1λT.

5.2.3 Peptide synthesis and rabbit immunization

Peptides were synthesized by AnaSpec Inc. (San Jose, CA). The two peptides, NIHPW6 (sequence ELISEMIRHPYETQSDSFYF) and NIHPW7 (sequence MHNKADYSYSGTP), were prepared for rabbit injection as follows. 2 mg of each peptide were suspended in 1.5 mL of ddH$_2$O. To 750 μL of this, 30 μL of 4M NaCl and 750 μL of Freund's Complete Adjuvant (FCA; Sigma) were added. For subsequent injections, 1 mg of peptide was suspended in 1.5 mL of ddH$_2$O; To 750 μL of this, 30 μL 4M NaCl, and 750 μL of Freund's Incomplete Adjuvant (FIA; Sigma) were added.

Two rabbits were obtained for each peptide. Rabbits were maintained, injected, and bled by the Animal Services division of the Biological Sciences Department, University of Alberta. Rabbits 6K4 and 6K8 were used for NIHPW6; rabbits 6K5 and 6K7 were used for NIHPW7. The schedule of injections is shown in Table 5.1. For each bleed, 10 mL of blood was obtained except for the pre-boost bleed, which was only 2 mL. Blood was allowed to coagulate for 1 hr at room temperature; the serum was drawn off and stored at 4°C in the presence of 0.1% sodium azide as antibacterial agent.

5.2.4 Spot blot analysis

A Bio-Rad 96-well vacuum-assisted spot blot apparatus was used with Bio-Rad nylon membrane. All manipulations were performed at room temperature. Unless otherwise noted, solutions added to the membrane were adjusted to a volume of 200 μL with PBS. Peptides were dissolved in phosphate-buffered saline (PBS) and drawn through the appropriate slot for 15 min by gravity, and then drawn through by vacuum. The slots were washed 3× by drawing 100 μL of PBS (containing 0.1% tween) by vacuum. Approximately 100 μL of blocking buffer (3% BSA in PBS) was added to each slot and allowed to incubate for 30 min, after which the excess solution was drawn through by vacuum. After washing 3× with tween-PBS, the rabbit antisera were diluted appropriately with PBS and allowed to incubate on each slot for 30 min. After five rinses of tween-PBS, each slot was incubated for 30 min with alkaline-phosphatase conjugated mouse α-rabbit Ab (Boehringer Mannheim) diluted 1:10000 in PBS. Following three more tween-PBS rinses, the membrane was removed from the apparatus and allowed to incubate in the presence of BCIP/NBT substrate overnight (Harlow and Lane, 1988).

5.2.5 Fixation and antibody staining

Animals were prepared for antibody staining by three methods. "Freeze-cracking" was performed as follows. Worms were washed off of plates with EN buffer (10 mM EDTA, 100 mM NaCl, pH 7.4) into eppendorf tubes and pelleted by brief centrifugation. After two rinses the solution was removed to ~2× the volume of worms. 25 μL of the worm suspension was added to a slide pre-coated in poly-L-lysine. (These slides are prepared by spreadng ~40 μL of 0.1 mg/mL poly-L-lysine solution in ddH_2O, then baking at 37°C for 30 minutes.) A 22 × 50 coverslip was placed over the worms and the slide placed at -80°C for at least 15 min. The coverglass was quickly removed with a razor blade and the slides immersed in -20°C methanol for 4 min in a coplin jar. Slides were transferred through the following solutions: -20°C acetone for 4 min, then (at room temperature) 75% acetone, 50% acetone, 25% acetone, then Tween-TBS [Tween-TBS is 1 mL Tween 20 per L of Tris-buffered saline (TBS): 150 mM NaCl, 50 mM Tris pH 7.8].

Anti-HA antibodies (monoclonal 12CA5, Boehringer Mannheim) were diluted 1:100 in Tween-TBS + 1% BSA and added dropwise onto the slides (~50 µL per slide), which were then placed in a humidified chamber at room temperature for 3 hours. Slides were rinsed in Tween-TBS for one hour in a coplin jar and returned to the chamber. In the dark, the 2° Ab (TexasRed-conjugated goat α-mouse, Jackson Immunochemicals, diluted 1:250 in Tween-TBS-BSA) was added and allowed to incubate for at least one hour. The slides were rinsed again for one hour in Tween-TBS in a coplin jar. For microscopy, approximately 15 µL of 2% DABCO (Sigma) in 90% glycerol was added, and the worms covered with a 22 × 50 coverslip.

The other two immunohistochemical methods were adapted from published papers. For anti-serotonin staining, the paraformaldehyde fixation protocol described in Desai *et al.* (1988) was used. For anti-GABA staining, the paraformaldehyde/glutaraldehyde protocol was followed as described in Link *et al.* (1992). Preparation of animals for anti-GABA staining was confirmed by immunological detection of GABAergic neurons (data not shown). Following fixation and permeabilization of worms, the crude rabbit sera (pre-immune and 2nd test bleeds of animal 6K4) were diluted 1:200 in AbA buffer [(1% BSA, 0.5% Triton X-100, 0.05% sodium azide, 1 mM EDTA in PBS (50 mM Na_2HPO_4, 140 mM NaCl, pH 7.2)] and added to worms in 0.5 mL thin-walled microfuge tubes at 4°C overnight. Worms were rinsed 3× in 0.5 mL AbB (same as AbA, with only

0.1% BSA) at room temperature for at least 15 minutes per rinse. Worms were incubated for three hours at 37°C in the dark in the presence of Cy3-conjugated goat α-rabbit 2° antibody (Jackson Immunochemicals) diluted 1:200 in AbA. Finally, animals were rinsed 3× in AbB. For mounting, 7 µL of worm suspension (in AbB) was added to 7 µL of DABCO solution (described above) and covered with a 22 × 22 coverslip. For anti-HA detection, a monoclonal anti-HA antibody was used at 1:200 dilution in AbA (BAbCO Berkeley Antibody Company). The 2° Ab was a Cy3-conjugated goat α-mouse (Jackson Immunochemicals) diluted 1:200 in AbA.

5.2.6 Construction of HA-tagged unc-119

The plasmid pUR#91, a gift from W. Wadsworth, contains a 128-bp *Bgl*II fragment encoding a triplet hemagglutinin (HA) tag. The construction of plasmid pDP#MM016M, which contains a novel *Bam*HI site at the end of the UNC-119 coding region, was described in §4.2.1. The plasmid pDP#MM016B, which contains a novel *Bgl* II site just after the start of the UNC-119 coding region, was generated from pDP#MM016 by *in vitro* mutagenesis (see §4.2.1) using the primer MMA3 (5'-GGC AGA TCT ACA ACA ACA ATC G-3').

The *Bgl*II insertion from pUR#91 was cloned into the *Bam*HI site of pDP#MM016B to make pUHA1, which encodes UNC-119 containing a triplet HA tag at the amino terminus (HA::UNC-119). Similarly, the HA tag was cloned at the *Bam*HI site in pDP#MM016M to make pUHA2, encoding a carboxy-tagged UNC-119::HA. Plasmids were injected individually at a concentration of 100 ng/µL (Mello, 1991). Plasmids pUHA1 and pUHA2 have also been referred to as pDP#MM016HA and pDP#MM016HA2, respectively.

A fragment of pUHA1 containing the HA tag was cloned into pBluescript KS- and sequenced using Sequenase 2.0 (Amersham) to confirm the correct insertion of the tag. Plasmids were injected at a concentration of 100 ng/µL with an equal molar amount of pRF4, which encodes the dominant *rol-6* transformation marker (Mello, 1991).

5.2.7 Construction of heat-shock unc-119 fusions

The structural features of the heat-shock fusions are shown in Figure 5.2. The expression vector pPD49.78 (*hsp 16-2*; obtained from Colin Thacker) was used to construct heat-shock fusions. The *unc-119* cDNA2 insert was removed with *Bam*HI and *Kpn*I, which digest within the pBluescript polylinker, and cloned into the same sites in pPD49.78. This clone contains the complete 3' end of the *unc-119* cDNA, but begins 34 bp upstream of the start of exon III (see Chapter 2). pUHS1 is predicted to express the carboxyl 199 aa of UNC-119. pUHS2 was similarly constructed, but used the complete cDNA3 insert (liberated with *Bam*HI and *Kpn*I), and is predicted to encode the entire UNC-119 protein.

pUHS3, which encodes an UNC-119::GFP fusion, was constructed by isolating a 740-bp *Hin*DIII-*Nhe*I fragment of pUHS2, which includes the heat-inducible promoter fused to the *unc-119* cDNA up to the middle of exon III. Plasmid pUGF7 (see §4.7.2) was digested with *Hin*DIII and *Nhe*I, and the larger 5.6 kb backbone was ligated to the pUHS2 fragment to create pUHS3.

pUHS4 encodes an UNC-119::GFP fusion that includes that modified S65C GFP and the 3'UTR from *unc-119*. A 740-bp *Hin*DIII-*Nhe*I fragment of pUHS3 was isolated. Plasmid pUGF16 (see §4.2.1) was similarly digested and the 5.6 kb backbone was ligated to the pUHS3 fragment to make pUHS4.

5.2.8 unc-6::UNC-119 transgene

A 1.5-kb fragment containing the promoter from the *unc-6* gene was isolated from pIM#155 (a gift from W. Wadsworth) digested with *Eco*RI and *Xho*I.

This fragment was cloned into pBluescript KS(-) to make pDP#MM060. The complete *unc-119* cDNA was liberated from cDNA3 by the enzyme *Eco*RI and cloned into the same site in pDP#MM061. The *unc-6*::UNC-119 plasmid was introduced into *unc-119(e2498)* by germline microinjection (Mello *et al.*, 1991) at a concentration of 100 ng/µL. No coinjection marker was used.

5.2.9 Construction of pUBP1 yeast two-hybrid "bait" plasmid

An UNC-119 "bait" plasmid, shown in Figure 5.3, was constructed using the vector pAS1-CYH2 (a gift from Steve Elledge, Baylor College, Houston). First, a modified *unc-119* cDNA was used (cDNA6B) in which two novel *Bam*HI sites had been introduced by *in vitro* mutagenesis. The first site was introduced immediately downstream of the start codon using the primer MMA3 (see §5.2.6); the second was introduced just short of the stop codon by MMA4 (§5.2.2). The resultant *Bgl*II-*Bam*HI fragment retains the original *unc-119* start codon and is predicted to encode almost all of UNC-119.

As the upstream *Bgl*II site was out of frame with the cloning site in pAS1-CYH2, a short linker was introduced to bring the UNC-119 coding region into frame. The *unc-119* expression plasmid pDP#MM065 (§4.7.4) contains *Bgl*II and *Bam*HI sites separated by 37 base pairs, free of stop codons. The pDP#MM065 plasmid was digested with *Bam*HI, and the *Bgl*II-*Bam*HI insert was inserted at this site. The resultand lefthand cloning junction has the sequence 5'-GGATCT-3' and cannot be cleaved by *Bam*HI or *Bgl*II. Redigestion of this intermediate plasmid (pDP#MM141) with *Bgl*II and *Bam*HI releases a slightly larger fragment containing an upstream linker allowing the fusion junction to be ligated in the correct reading frame (plasmid pDP#MM143, renamed pUBP1).

Plasmid pUBP1 was packaged into single strands (an f1 origin is retained in the plasmid backbone) and sequenced using Sequenase 2.0 (Amersham) and the GAL4 primer MMA31 (5'-ATG CCT CTA ACA TTG AGA C-3'), both with dGTP and dITP reaction mixes according to the manufacturer's instructions; the lefthand junction sequence (shown in Figure 5.3) was found to be free of errors.

5.2.10 Yeast two-hybrid screen

The following work with the *unc-119* bait plasmid was performed by Gary Ritzel, University of Alberta. The bait plasmid pUBP1 was transformed into the *S. cerevisiae* strain PJ69-4A (James *et al.*, 1996), and selected on SC (synthetic complete) -trp media. Approximately 2µg of the *C. elegans* cDNA library DNA were transformed into PJ69-4A using an RNA carrier-based protocol at an efficiency of approximately 10^4 cfu/µg. Transformants were selected on two 150mm SC -leu -his -trp -adenine + 1 mM 3AT (3-amino-1,2,4-triazole) plates and allowed to grow for two days (G. Ritzel, pers. comm.). Approximately 50 transformants were isolated and patched to fresh plates. The inserts belonging to the library plasmid of 15 of the strongest colonies were PCR-amplified using the vector specific primers ADH-T7 and ADH Term, and sequenced (G. Ritzel, pers. comm.). In preparation for mating confirmation against the *S. cerevisiae* Y187 bait tester strains, small-scale plasmid preparations of the library plasmids were cotransformed with pUBP1 into strain Y190 under histidine selection; all 15 transformant strains grew well. Activation of the reporter HIS marker is required for growth, suggesting that in Y190, pUBP1 and the library clones can activate GAL-UAS transcription of the HIS3 reporter (G. Ritzel, pers. comm.). The definitive confirmatory experiments were not performed.

5.3 Results

5.3.1 *Expression of recombinant UNC-119 and HRG4 in E. coli*

Antibodies can be most practically raised by injecting purified protein into a suitable vertebrate host such as rabbit or mouse, in order to generate polyclonal antisera (Harlow and Lane, 1988). The main difficulties are the antigenic potential of the hapten and its purity. Consideration of the amino acid sequence of the protein of interest can be used to gauge which portions are likely to be antigenic. For the second problem, parts (or all) of the protein can be produced as fusions to a carrier protein, such as GST (glutathione S-transferase), in *E. coli* and affinity purified on a large scale. Prior to injection, the recombinant protein can be further treated with thrombin to remove the GST carrier.

Several plasmids were constructed using the pGEX-1λT vector (Pharmacia) to allow bacterial expression of GST fusions to various portions of *C. elegans* UNC-119, *C. briggsae* UNC-119, and human HRG4. A summary of the portions of the coding regions used in these vectors is shown in Figure 5.1. Fusions of GST to certain portions of UNC-119, rather than the entire coding region, were made in order to anticipate problems that can occur with full-length heterologous proteins in *E. coli*, such as formation of inclusion bodies, protein instability, and toxicity.

A small number of attempts were made to demonstrate expression of the GST fusion proteins in *E. coli* upon induction with isopropyl thiogalactoside (IPTG). However, derivatives of the *E. coli* strain BL21(DE3), obtained by transformation with the *C. elegans* vectors pUFP1 to pUFP4, showed low expression upon induction. The GST::HRG4 fusion pUFP5 did show overexpression, however; the reason for this disparity is unknown. To rule out mutation in the plasmid backbone, the HRG4 cDNA in pUFP5 was excised and replaced with a similar cDNA insert as pUFP2, to create pUFP6; however, no difference in expression of pUFP6 over pUFP2 was noted (data not shown). These experiments were repeated with similar results, and the GST fusion with *C. briggsae* UNC-119, pUFP7, has subsequently been shown to express weakly in *E. coli* (S. McGuire, pers. comm.).

A single experiment to purify the pUFP1-encoded GST::UNC-119 fusion protein, expressed in *E. coli* BL21(DE3) yielded a weak band of the expected size (~41 kD) along with a strong band of approximately 26 kD. This size corresponds to native GST, suggesting that the fusion protein may have been degraded during purification. The reason for the poor recovery of GST::UNC-119 was not determined, and no further purifications of recombinant UNC-119 were attempted.

5.3.2 Testing of antisera to two HRG4 epitopes

An opportunity arose to obtain two synthetic peptides to use for the generation of antibodies. Two peptides, NIHPW6 and NIHPW7, were designed based on segments of the HRG4 coding region that were predicted by computer algorithm to be suitable epitopes (Paul Wong, pers. comm.). These segments, and their relative conservation with UNC-119, are shown in Figure 5.4. The peptides were injected into rabbits (two per peptide) at a standard concentration and at standard intervals (Harlow and Lane, 1988); test bleeds were obtained appropriately (Table 5.1).

A "spot blot" approach was used to test the pre-immune and second bleeds from rabbits 6K5 and 6K7 (putative antisera to peptide NIHPW7). A two-dimensional array was chosen to titre any possible cross-reaction. Along one axis, peptides were spotted on in amounts of 1 ng, 10 ng, 100 ng and 1000 ng.

Along the second axis, crude rabbit antisera were incubated on the spots as 1 µL, 0.1 µL, 0.01 µL and 0.002 µL. As a positive control, 1 µL of each rabbit serum was spotted directly onto the membrane prior to blocking, and as a negative control, 200 µL of buffer were spotted. For the these controls, 0.1 µL of serum from rabbit 6K5 was used as the primary antibody.

Following an overnight incubation in substrate, however, no cross-reaction was observed for the sera against either peptide. The positive controls showed moderate cross-reaction by the secondary antibody, confirming that the detection method was not at fault. This experiment was subsequently repeated with different reagents and rabbit bleeds (H. Lemon and P. Wong, pers. comm.) with similar negative results. Furthermore, no cross-creaction was obtained when the secondary bleeds from any of the four rabbits were used on human retina tissue (D. Swanson and D. Valle, pers. comm.).

A second independent method was chosen to test the rabbit sera for putative cross-reaction against recombinant HRG4 protein. The ability of an *unc-119*::HRG4 transgene to function in place of *unc-119* in *C. elegans* has already been described, as has the neural localization of a tripartite *unc-119*::HRG4::GFP fusion (Chapter 4). These results were consistent with expression of HRG4 in *C. elegans*. Therefore, antibodies to HRG4 epitopes are likely to recognize the recombinant protein in transgenic nematodes carrying an *unc-119*::HRG4 transgene.

Crude serum from one of the NIHPW6-immunized rabbits (6K4) was used in a preliminary attempt to visualize HRG4 protein in the *C. elegans* strain DP142. This strain is homozygous for the *unc-119* null mutation *ed3* and carries the *unc-119*::HRG4 plasmid pDP#MM103 as a chromosomally integrated array. DP142 animals were fixed using methods for the detection of the neurotransmitters GABA (γ-aminobutyric acid) or serotonin (Desai *et al.*, 1988; Link *et al.*, 1992).

A summary of the results is compiled in Table 5.2. There was some neural staining in the pharynx which was not evident in animals stained with pre-immune serum (Figure 5.5A). However, strong or obvious cross-reaction, particularly neural, was not observed. It is not known whether the staining seen was genuine HRG4 cross-reaction or artifact; if it was genuine, pre-incubation of the serum with purified HRG4 peptide would be expected to block the cross-reaction. This experiment was not performed, nor were other methods of fixation attempted.

5.3.3 Immunological detection of hemagglutinin (HA) tagged UNC-119

An alternative to immunological detection of a protein itself is the addition of a known epitope through insertion into the protein coding region (reviewed in Kolodziej and Young, 1991). Commerically available antisera can then be used for indirect whole-mount immunofluorescence. Theoretically, the small size of such a tag also minimizes the likelihood of its interference with the true native expression pattern of the gene.

One epitope that has been used for this purpose is the hemagglutinin epitope (HA) from the influenza hemagglutinin protein (Wilson *et al.*, 1984). A *C. elegans* transgene encoding a triplet HA tag at the UNC-6 amino terminal coding region allowed the identification of neurons expressing HA::UNC-6, using anti-HA antibodies (Wadsworth *et al.*, 1996). Using a DNA fragment encoding the same triplet tag, both amino-tagged and carboxyl-tagged versions of *unc-119* were constructed. Germline transformation of *unc-119* mutants with either pUHA1 (encoding HA::UNC-119) or pUHA2 (UNC-119::HA) was found to produce phenotypically rescued progeny, consistent with the function of UNC-119 carrying the small HA tags. Therefore, antibodies to HA should reveal the *in vivo* localization of the fusion proteins.

Transgenic *unc-119(e2498)* mutants, rescued by the pUHA1 transgene, and the *unc-6*::HA strain *urls1* (a gift from W. Wadsworth), were fixed and permeabilized using a standard methanol/acetone protocol (§5.2.5). Animals were stained with a monoclonal anti-HA antibody and a secondary TexasRed conjugated secondary antibody. However, despite several attempts, good staining was not obtained with this method except on one occasion, which showed apparent embryonic neural staining, and some neural cell bodies in adults (Figure 5.5, panels B-D). The *urls1* control failed to stain, however. A control experiment using a monoclonal anti-βgal antibody did give good neural staining of *unc-119*::lacZ transgenic animals (data not shown), suggesting the fixation method was not at fault.

A second attempt was made using a second anti-HA antibody with the fixation and permeabilization protocols optimized for detection of GABA and serotonin (described above). Similarly, very little staining was observed for the HA::UNC-119 and UNC-119::HA strains (Figure 5.5, panels F and G). Some neural staining was seen in the HA::UNC-6 positive control (Figure 5.5E). The main reason for the poor antibody detection seems to have been the fixation method used (see Discussion). These experiments were abandoned, in favour of the simpler functional GFP fusion approach (Chapter 4).

5.3.4 Expression of heat-shock driven unc-119 expression

Expression of *unc-119* reporter transgenes demonstrates that the *unc-119* message is expressed beginning in the early embryo, and is maintained through adulthood, in the nervous system (Chapter 2). It is possible, however, that the gene product is not actually required in the adult. As the mutant phenotype suggests that UNC-119 has a role in axon guidance, it would seem unnecessary to maintain its expression once neurogenesis is complete, although a maintenance role cannot be ruled out (see Chapter 6). As there are no conditional alleles of *unc-119* that would allow experimental removal of gene function in the adult, an attempt was made to rescue *unc-119* by heat-shock expression.

In Chapter 2, it was shown that a 'minigene' expressing the *unc-119* cDNA can rescue the *unc-119* phenotype. Therefore, neural expression of the cDNA at the appropriate times in development would be expected to rescue. The heat-shock promoter *hsp16-2*, whose expression has been characterized (Stringham *et al.*, 1992), was fused to an incomplete *unc-119* cDNA (pUHS1) and a complete cDNA (pUHS2), as shown in Figure 5.2. Wild type (N2) animals were made transgenic for pUHS2 by germline transformation (Mello, 1991). Heterozygous *unc-119/+* males were crossed into Rol transformants, and Unc Rol progeny were segregated in the F_2. Mixed-stage plates of the *unc-119* transformants were incubated at 33°C for 15 minutes and returned to room temperature for one hour. Although this cycle was repeated several times in succession, there was no evidence of any rescued animals. This may have been the result of a lack of proper transgene expression, inappropriate induction, or constitutive requirement for UNC-119.

In order to determine whether the heat-shock transgene was actually being expressed, this experiment was modified to allow visualization of the heat-shock induced protein. It has been shown that carboxyl fusions of UNC-119 to GFP are able to rescue the *unc-119* phenotype (Chapter 4). A chimeric plasmid, consisting of part of pUHS2 and pUGF7 (a full-length UNC-119::GFP fusion to the original GFP cDNA), was constructed. This plasmid, pUHS3, is predicted to express a full-length UNC-119::GFP fusion under control of the heat-shock promoter (Figure 5.3). Wild-type animals were made transgenic for pUHS3 and pRF4 (*rol-6D*). A mixed-stage plate of transgenics was incubated at 33°C for three hours, and examined under fluorescence microscopy. No fluorescence was visible immediately following induction. As a several-hour lag time has been observed for unmodified GFP to become fluorescent in bacteria (Cormack *et al.*, 1996), it is possible that the UNC-119::GFP fusion protein was made but had not yet become fluorescent. Animals were observed 20 hours post-induction and indeed found to have accumulated GFP in several tissues, including those that do not express the *unc-119*::GFP transgene (Figure 5.6).

It can be concluded from these experiments that the heat-shock promoter *hsp16-2* is indeed capable of expressing UNC-119::GFP following heat shock. Within the pharynx, the fusion protein appears to accumulate in the cytoplasm, and is apparently stable for at least 20 hours after induction. It is likely, therefore, that expression of untagged UNC-119 was occurring in animals transgenic for pUHS2, but the treatments were insufficient to rescue the *unc-119* phenotype.

A second UNC-119::GFP heat-shock plasmid, pUHS4 (see Figure 5.3), was made to the S65C variant of GFP to see if visualization of the heat-shock induced protein fusion could be improved. N2 transgenics were made with pUHS4, but these animals were not examined. No further attempts were made to rescue *unc-119* by this approach.

5.3.5 Attempted rescue by an unc-6::UNC-119 transgene

An alternative to using a heat-shock::UNC-119 transgene to rescue *unc-119* is to rescue by expression of UNC-119 under the control of a second, well-characterized promoter. The neural gene *unc-6* is expressed in pioneering neurons and glia of the ventral midline (Wadsworth *et al.*, 1996), and therefore expresses in a subset of the cells that drive *unc-119* transgene reporters (Chapters 2,3,4).

An *unc-6*::UNC-119 fusion gene was constructed, and introduced into *unc-119* (*e2498*) mutants by germline transformation (Mello *et al.*, 1991). One day after injection, at least five rescued L1 animals were seen, distributed between two independently injected parents. Rescue was assessed solely on the qualitative observation of locomotion. These rescued progeny were transferred to separate plates. After four days, however, they were strongly uncoordinated in a manner similar to *unc-119* mutants. None of these animals gave rise to rescued progeny, likely as a result of loss of the extrachromosomal array (Mello *et al.*, 1991).

The apparent conclusion is that *unc-6*-directed expression of UNC-119 can rescue the locomotory defect until hatching, but subsequent to this stage, this rescue is lost. This is superficially consistent with the largely embryonic expression of an epitope-tagged version of *unc-6* (Wadsworth *et al.*, 1996), and seems to imply that, for the locomotory defect, expression of *unc-119* is required beyond L1. Alternatively, it may be that embryonic neurons required for locomotion at hatching are properly guided, but those that are derived post-embryonically are not. Among the motor neurons (§1.8.3), *unc-6*::HA expression only occurs in the postembryonically derived VA2-VA12 and VB3-VB11 neurons, providing a supplemental source of the UNC-6/netrin gradient (Wadsworth *et al.*, 1996). Therefore, postembryonic expression of UNC-119 in these neurons may not be sufficient to rescue.

5.3.6 Screening of a yeast two-hybrid library

The "two-hybrid" system in the yeast *Saccharomyces cerevisiae* is a powerful method of identifying putative protein-protein interactions with a gene product from a cloned cDNA of interest (Fields and Song, 1989; Chien *et al.*, 1991). An attempt was made using a modified version of this system which was found to reduce the amount of false positives (James *et al.*, 1996). An *unc-119* "bait" plasmid, pUBP1, was constructed (§5.2.9 and Figure 5.3). Yeast of strain PJ69-4A were transformed sequentially, first with pUBP1. The transformed strain was unable to cause false marker expression, a prerequisite for subsequent use in the screen (G. Ritzel, pers. comm.).

Second, a plasmid library of cDNAs cloned downstream of the GAL4 activation domain was transformed into the strain. Yeast cells in which a cDNA clone encodes an UNC-119-interacting protein may allow association of the two GAL4 fusion proteins, activating expression of three reporter genes present chromosomally in PJ69-4A. This allows for direct selection of putative interaction clones on appropriate media (James *et al.*, 1996).

Approximately 50 positives were recovered and replated; some of the best-growing colonies were chosen for further analysis. The sequence of six of the cDNAs selected in this screen have all been sequenced independently by the *C. elegans* genome sequencing consortium (Wilson *et al.*, 1994). A summary is presented in Table 5.4. None of the clones have been associated with known *C. elegans* mutations.

Unfortunately, it is difficult to draw any firm conclusions from these data, as there are several artifacts that can be expected to cause false positives. First, expression of an out-of-frame cDNA may produce an interacting polypeptide. Second, real interactions may be demonstrated for proteins that may never be coexpressed in *C. elegans* both temporally and spatially, and are hence not biologically relevant. Third, and most importantly, the positives were not rigorously tested for false reporter activation (i.e. non-UNC-119-specific).

Given these caveats, the recovery of two different cDNAs corresponding to the same gene, EF-1α (Elongation Factor 1α), is interesting because of the properties of this protein (see Discussion).

5.4 Discussion

5.4.1 Comments regarding this section

In this chapter, several pilot projects were initiated with the long-term goal of understanding the *in vivo* function of the UNC-119 protein. Most of these yielded, at best, preliminary data; therefore, any conclusions must necessarily be understood to be working hypotheses based on unrepeated experiments. Some of this work has been taken over by subsequent members of the laboratory. Each set of results will be discussed, followed by a set of concluding remarks.

5.4.2 Expression of recombinant GST::UNC-119

Several GST fusions were made in an attempt to express recombinant UNC-119 in *E. coli* for large-scale purification. Of fusions made to full-length *C. elegans* UNC-119, *C. briggsae* CbUNC-119, or human HRG4, only the GST::HRG4 fusion showed strong induction in bacteria. Both HRG4 and UNC-119 have extended regions of similarity, and the same HRG4 cDNA used in the GST fusion can fully replace *unc-119* in *C. elegans*, suggesting that UNC-119 and HRG4 have similar biochemical properties (Chapter 4). Presumably, the basis for poor expression lies in the unconserved residues found in UNC-119. This is supported, weakly, by the failure of CbUNC-119 (the *C. briggsae* homologue) to overexpress: these proteins share 90% identity across their entire lengths (Chapter 3).

Weak expression need not preclude purification of the recombinant protein on a large scale. There was some evidence, in a preliminary purification attempt, that some recombinant full-length GST::UNC-119 could be obtained, although the amount was not quantified. Furthermore, small amounts of recombinant protein can be sufficient for an immune response (Harlow and Lane, 1988). Alternatively, recombinant HRG4 may be used as an immunogen. The polyclonal antisera so obtained may include UNC-119 cross-reacting antibodies and allow detection of the polypeptide, or alternatively, HRG4 itself could be detected in strains carrying the integrated *unc-119*::HRG4 transgene *edIs16* (Chapter 4). In addition, the existence of a partial cDNA for the *Drosophila* DmUNC119 homologue, bounded by *Bam*HI sites, would allow its cloning into pGEX1λT for the same set of experiments.

5.4.3 Antibodies to HRG4 synthetic peptides

An attempt to generate antibodies to HRG4 was made using synthetic peptides having the same sequence as two segments near the HRG4 carboxy terminus (Figure 5.4). Of the two peptides synthesized, NIHPW7 would seem to be the more likely to generate antibodies that can cross-react to native UNC-119, as it shares nine identical contiguous residues. Unfortunately, significant cross-reaction to the peptides was not observed on spot blots. The conclusion appears to be that the epitopes may not be sufficiently immunogenic.

Under the supposition that subtle changes might have occured to the epitope *in vivo* that precluded efficient antibody binding on membranes, rabbit serum was tested on fixed animals expressing the *unc-119*::HRG4 fusion. It was interesting that some weak signal was observed. This may have been genuine HRG4 cross-reaction, as animals treated with pre-bleed serum did not show this staining. The expression of a functional HRG4::GFP fusion suggests that the *unc-119*::HRG4 transgene directs similar expression of HRG4 (Chapter 4); in turn, there should be strong anti-HRG4 signal throughout the nervous system in transgenic animals. If the rabbit serum can indeed cross-hybridize to recombinant HRG4 expressed in *C. elegans*, it may be that an alternative fixation protocol may allow detection. The procedure that allows detection of the neurotransmitter GABA requires long incubation times in glutaraldehyde-paraformaldehyde solution (Link, 1992); these chemicals can destroy sensitive epitopes and inhibit antibody recognition (Harlow and Lane, 1988).

5.4.4 Hemagglutinin tagged UNC-119

The expression of HA-tagged versions of *unc-119* demonstrated that the addition of the epitope did not interfere with UNC-119 function, as these constructs also rescued the *unc-119* locomotory defect. Attempts to detect the recombinant HA triplet tag met with mixed success. On one occasion, expression was observed in embryos, suggesting that the epitope can be detected by anti-HA antibodies, but that the methodology itself was not optimized. A subsequent attempt using the fixation protocols for the neurotransmitters GABA and serotonin (Desai *et al.*, 1988; Link *et al.*, 1992) may have destroyed the epitope, which could account for the failure to detect strong expression of HA-tagged UNC-119. This was consistent with failure to obtain a strong signal in HA::UNC-6 control animals.

These experiments were not followed up, largely due to the consideration of the initial reasons for epitope tagging. As GFP fusion vectors became available to the *C. elegans* community, it became apparent that many complete GFP fusion transgenes can be functional, while retaining fluorescence. In these cases, it is simpler to observe GFP fluorescence than to visualize an epitope tag by indirect immunofluorescence, especially since animals can be observed *in* *vivo* or with minimal, non-destructive preparation (Chalfie *et al.*, 1994). The use of GFP instead of an epitope still suffers from the same caveats as epitope tagging, such as interference with protein function, post-translational destruction of the tag, or phenotypic effect caused by multicopy transgene arrays. The only difference is the size of the added polypeptides: the triplet HA tag only adds ~40 aa to a fusion protein, while GFP adds approximately 270 aa (§4.10). However, verification of simultaneous fluorescence and function of a fusion transgene in *C. elegans* is straightforward.

5.4.5 Heat-shock expression of UNC-119

Attempts to rescue the *unc-119* phenotype by tissue non-specific expression under the control of the *hsp16-2* promoter (Stringham *et al.*, 1992) did not meet with success. If UNC-119 is required constitutively, heat-shock treatment may have to be applied at regular intervals throughout embryogenesis and larval stages, to achieve rescue. However, the apparent accumulation of GFP signal from the *hsp16-2*::UNC-119::GFP fusion pUHS3 in pharynx suggests that the fusion protein may be stable for at least 20 hours (Figure 5.6), although this does not necessarily mean that UNC-119 function is retained.

It is formally possible, however, that UNC-119 may be turned over rapidly in neurons. This is somewhat consistent with the failure to observe high degrees of fluorescence from the pUHS3 transgene in neurons 20 hours following induction; however, somatic mosaicism of the transgene or subthreshold GFP accumulation (i.e. not sufficient for a fluorescence microscope signal) cannot be ruled out.

This experiment could be repeated by adapting a thermal cycler to apply the necessary incubations over a longer time period. Furthermore, the heat-shock transgenes could be integrated, as was done for several *unc-119* fusion transgenes (Chapter 4), to eliminate somatic mosaicism. Alternatively, a conditional allele would be the most convenient method to ascertain requirement for UNC-119 function (discussed in Chapter 6).

5.4.6 Partial rescue by an unc-6::UNC-119 transgene

The ability of an *unc-6*::UNC-119 transgene to confer partial rescue to an *unc-119* mutant is good confirmation that *unc-119* itself expresses in neurons, as the *unc-6* gene is known to express in a subset of the nervous system (Wadsworth *et al.*, 1996). The loss of rescue through the larval stages seems to imply that UNC-119 function is required in the adult, or that this function is required in *unc-6* non-expressing cells. As there are better ways to address this question, this line of investigation does not seem to bear further study.

However, a properly controlled version of this experiment would require coinjection of the *unc-6*::UNC-119 plasmid with a transformation marker, such as pRF4 (*rol-6D*; Mello *et al.*, 1991), or the *unc-119*::GFP marker pUGF12 (Chapter 4), to follow the transgene. Alternatively, the *unc-6* promoter could be used to drive expression of UNC-119::GFP [analogous to pUGF16 (Chapter 4) or pUHS3 (see above)]. Furthermore, the transgene should also be chromosomally integrated to eliminate somatic mosaicism (Chapter 4).

5.4.7 Preliminary yeast two-hybrid screen with UNC-119

A yeast two-hybrid library was screened in an attempt to identify gene products that interact with UNC-119. In the preliminary screen described here, several positives were recovered, although these were not subjected to rigorous testing (G. Ritzel, pers. comm.). The yeast strain used for the screen, PJ69-4A, has been described as reducing (or eliminating) the occurrence of false positives inherent in earlier approaches, while increasing sensitivity to weaker interactions

(James *et al.*, 1996). This suggests that the identified cDNAs are more likely to represent real interactions. However, a control screen, using a bait prone to false positives for example, was not performed.

Nonetheless, the recovery of two different cDNAs for the *C. elegans* homologue of EF-1α suggests that this protein may interact with UNC-119. This particular gene provokes interesting models for UNC-119 function. EF-1α is the most abundant component of the protein synthesis machinery, interacts with the cytoskeleton, and bundles actin filaments and microtubules (reviewed in Condeelis, 1995). Depending on the molar ratio of EF-1α to actin, this protein can either cause the polymerization or depolymerization of actin filaments (Murray *et al.*, 1996). It is known that 'dynamic instability' of microtubules is required for growth cone extension, as chemicals that interfere with microtubule polymerization affect growth cone motility (Williamson *et al.*, 1996). As a growth cone defect is consistent with the observed *unc-119* cellular defect (Chapter 4), an indirect or accessory role for UNC-119 in the interaction of actin and microtubules could be proposed. A plausible model is described in Chapter 6.

Biochemical verification of this interaction would be the first necessary step if this hypothesis is to be given experimental credibility. Fortuitously, antibodies to EF-1α exist in many systems (e.g. Dharmawardhane *et al.*, 1991), and given the high degree of conservation between cross-species homologues, confirmation of this suspected interaction may require no new resources. In particular, *C. elegans* strains rescued by a HA-tagged version of *unc-119* could be used to check for an *in vivo* interaction by coimmunoprecipitation with anti-HA antibodies, followed by polyacrylamide gel electrophoresis, Western blot, and immunodetection of EF-1α (Harlow and Lane, 1988).

A further control would be to show that a GAL4::HRG4 fusion bait plasmid could activate reporter expression from a GAL4(BD)::EF-1α fusion, as the ability of HRG4 to functionally replace UNC-119 suggests these proteins should share the same interactions *in vivo*. In addition, such a 'directed' two-hybrid analysis could be used to search for potential interactions between UNC-119 and other proteins known to affect axon outgrowth and fasciculation, such as UNC-33, UNC-44, and UNC-73 (McIntire *et al.*, 1992).

Often, mutation in the gene encoding an interacting protein can produce a similar phenotype. For example, mutations in *unc-51* and *unc-14* have similar phenotypes, and their gene products have been shown to interact (Ogura *et al.*, 1997). It is unclear whether inactivation of EF-1α will even result in viable animals, as this protein plays a role in protein syntehsis (Condeelis, 1995). Consequently, a reverse genetics approach such as an EMS deletion screen (Jansen *et al.* 1997), or RNA-induced phenocopy (Mello *et al.*, pers. comm.), may not give a clear answer. If the interaction between UNC-119 and EF-1α is shown to be real, it would be the first evidence to implicate a role for UNC-119 in assembly of the cytoskeleton.

5.4.8 Concluding remarks

This set of experiments described allows an assignment of priorities in assessing UNC-119 requirement and function. In particular, the most important experiments to pursue are the generation of antibodies to UNC-119 and the repetition of the yeast two-hybrid screen. These will provide useful resources for future experiments, and give results bearing directly on UNC-119 function. For example, anti-UNC-119 antibodies can be used to verify interactions predicted by yeast two-hybrid data, and for immunolocalization of UNC-119 protein. It is hoped that these pilot experiments will expedite the analysis of the UNC-119 protein for future investigators.

Acknowledgement

I gratefully acknowledge the assistance of Paul Wong, who obtained the HRG4 synthetic peptides and provided helpful advice; Gary Ritzel for performing the yeast two-hybrid screen, sequencing the positives and performing database searches; Paul Stothard for assistance with construction of the bait plasmid; Petra Jäckle-Baldwin for providing prepared pGEX1λT plasmid and for help with protein gels; William Wadsworth for providing the pUR#91 and pIM#155 plasmids containing the triplet HA tag and *unc-6*, respecitvely, and the *urls1* strain; Colin Thacker for sending the heat-shock promoter plasmid; Paul Wong and Mark Peppler for providing laboratory space and materials for the testing of the rabbit antisera; Shawna McGuire for testing expression of the GST fusions in *E. coli*; Gina Broitman for help with the first set of spot blot tests with anti-HRG4 antisera, and for providing books and resources pertaining to immunizations and analysis of antibodies; Heather Lemon and Paul Wong for repeating the analysis of the HRG4 antisera; Deb Swanson for testing the rabbit antisera against human retina; and Gian Garriga for suggesting a directed yeast two-hybrid approach with UNC-119.

Date	Description
23 Oct 96	pre-immune bleed (10 mL)
23 Oct 96	1st injection (1 mg peptide)
20 Nov 96	pre-boost bleed (2 mL)
20 Nov 96	2nd injection (0.5 mg peptide)
29 Nov 96	1st test bleed (10 mL)
18 Dec 96	3rd injection (0.5 mg peptide)
8 Jan 97	2nd test bleed (10 mL)
4 Mar 97	4th injection (0.5 mg peptide)

Table 5.1 Rabbit injections with synthetic HRG4 peptides, and test bleeds.

Strain	Fixation	Serum	Notes
DP142	P + G	6K4 pre	• some surface antigens (speckled) • cross-reaction near embryos in centre of hermaphrodites • tip of mouth (entry of 2° Ab?)
DP142	P + G	6K4 2nd	• staining in pharyngeal bulb, apparently a cell body and process • vulva, tip of head
DP142	P	6K4 pre	• similar to P + G fixed animals
DP142	P	6K4 2nd	• pharyngeal neurons, paired cell bodies • some staining in periphery of proximal gonad • coelomocyte

Table 5.2 Summary of staining seen in strain DP142 with rabbit antisera to HRG4 synthetic peptide NIHPW6 (see Figure 5.4). Strain DP142 is homozygous for *unc-119*(*ed3*) and carries the *unc-119*::HRG4 transgene as a chromosomally integrated array. P, paraformaldehyde; G, glutaraldehyde; pre, pre-immune serum; 2nd, second test bleed.

Strain	Fixation	Antibody	Notes
urIs1	M + A	Boehringer mAb	• No good staining observed
e2498; [pUHA1]	M + A	Boehringer mAb	• Some neural embryonic staining, similar to *unc-119*::GFP expression • Some cells in nerve ring and ventral nerve cord
urIs1	P + G	BAbCO mAb	• cell bodies in nerve ring and ventral nerve cord • often saw cell body in posterior region • some nonspecific speckling on outer surface
urIs1	P	BAbCO mAb	• similar to above, but slightly weaker • cells in nerve ring, ventral nerve cord
e2498; [pUHA1]	P + G	BAbCO mAb	• some faint cell bodies • ALM process • neurons in preanal ganglion • some nerve ring cells
e2498; [pUHA1]	P	BAbCO mAb	• faint cell bodies, one ALM process • occasional cell bodies in nerve ring

Table 5.3 Summary of anti-HA staining seen in *C. elegans* strains carrying HA fusion transgenes. The *e2498*; [pUHA1] animals are homozygous for *unc-119*(*e2498*) and are rescued by the *unc-119*::HA amino-terminal fusion plasmid pUHA1. Strain *urIs1* is a chromosomal integrant of an *unc-6*::HA fusion (Wadsworth *et al.*, 1996). M, methanol; A, acetone; P, paraformaldehyde; G, glutaraldehyde; mAb, monoclonal antibody.

Clone #	Cosmid	Identity of Protein
2	F13D12	amino acid synthesis enzyme
3	n/a	yeast GAL4
6	F01G10	transketolase-like
8	R03G5	Elongation Factor 1 α
22	F32H2	fatty acid synthetase-like
23	R03G5	Elongation Factor 1 α

Table 5.4 Identity of some of the cDNAs obtained in a preliminary yeast two-hybrid screen using the bait plasmid pUBP1 (Figure 5.3). "Cosmid" refers to the physical map location from the *C. elegans* genome sequencing project (Wilson *et al.*, 1994). The sequences and database search results were provided by Gary Ritzel, University of Alberta.

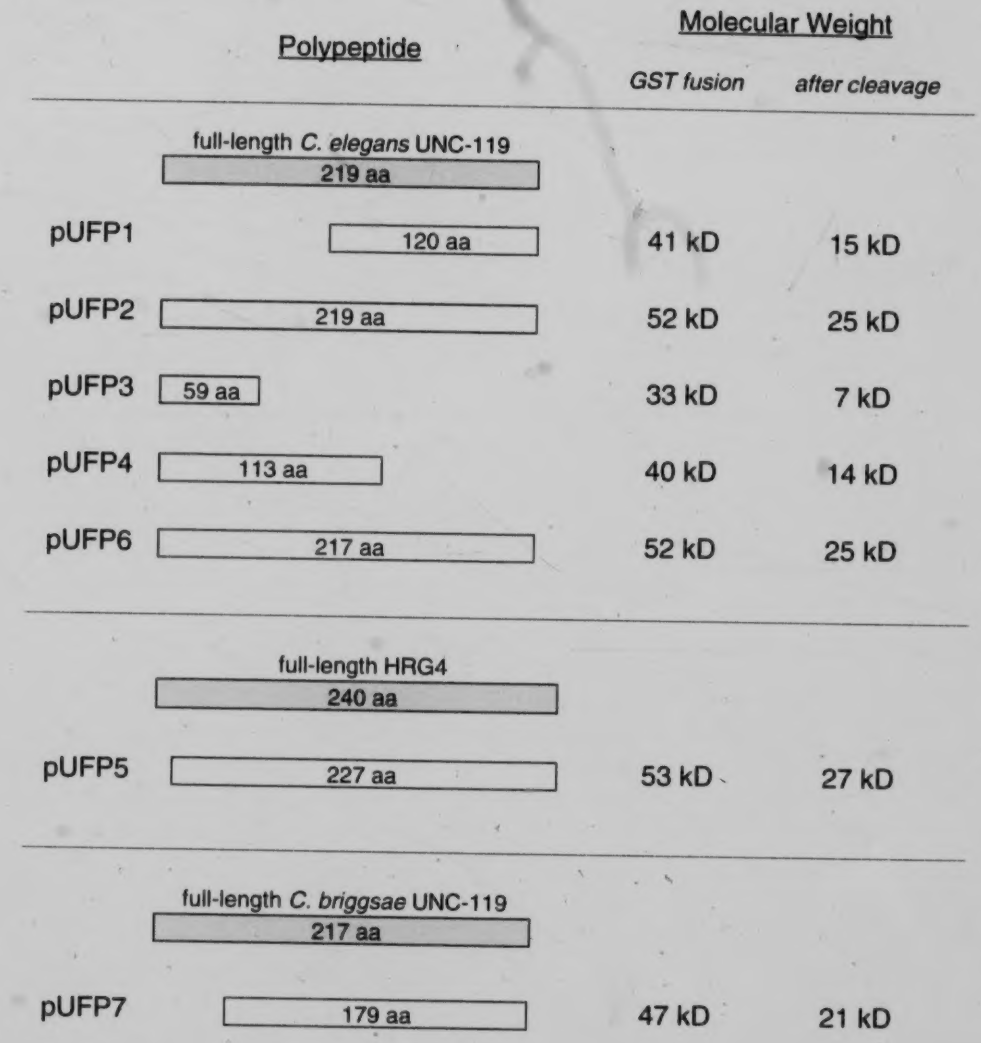
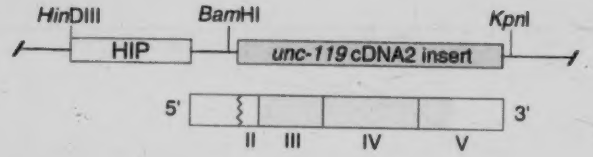
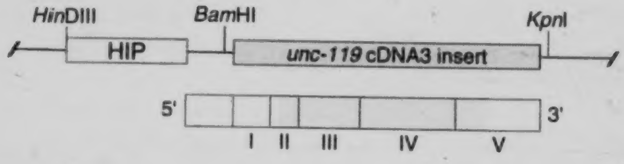


Figure 5.1 Schematic of GST fusion plasmids. The molecular weights of the fusion proteins were determined with DNA Strider 1.2 for the Macintosh (Marck, 1988). 'After cleavage' refers to the molecular weight of the polypeptide released after digestion of the GST fusion with thrombin (in addition to the ~26 kD GST portion).

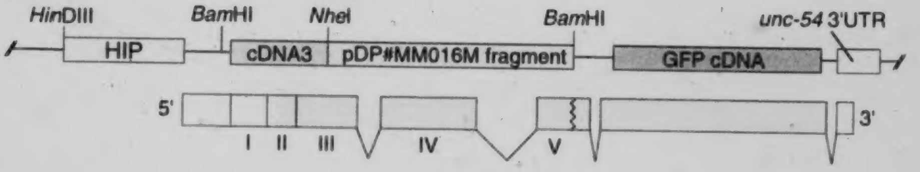
pUHS1 (HIP::UNC-119)



pUHS2 (HIP::UNC-119)



pUHS3 (HIP::UNC-119::GFP)



pUHS4 (HIP::UNC-119::GFP)

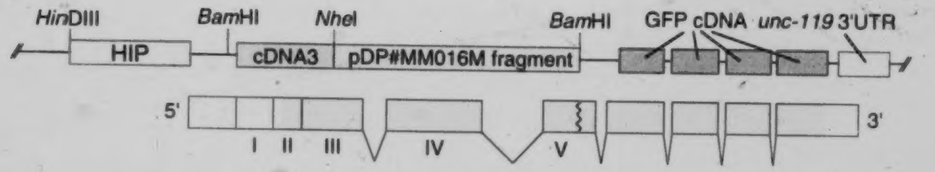


Figure 5.2 Schematic diagram of relevant structures in heat-shock fusion plasmids. The putative heat-shock induced mRNA is shown. Roman numerals indicate *unc-119* exons, and shaded regions depict the open reading frame (ORF). Jagged lines indicates that only part of a particular exon is included. The backbone plasmid for pUHS1-3 is pUC19 and that of pUHS4 is pBluescript KS(-). The sizes of the relative domains are only approximate and therefore not to scale. HIP, heat-inducible promoter *hsp16-2* (Stringham *et al.*, 1992)

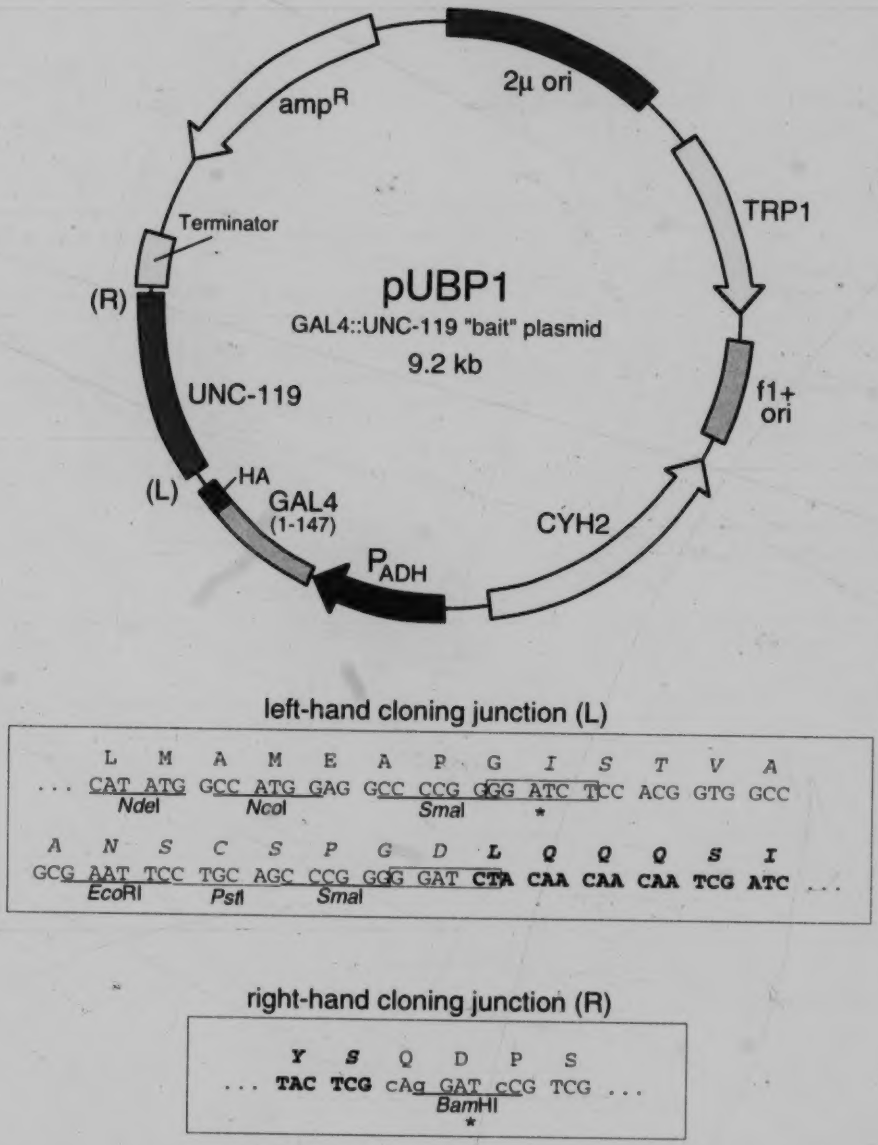

2μ ori
TRP1
f1+ ori
CYH2
P_ADH
GAL4 (1-147)
HA
(L)
UNC-119
(R)
Terminator
amp^R
pUBP1
GAL4::UNC-119 "bait" plasmid
9.2 kb
left-hand cloning junction (L)
L M A M E A P G I S T V A
... CAT ATG GCC ATG GAG GCC CCG GGG ATC TCC ACG GTG GCC
NdeI NcoI SmaI *
A N S C S P G D L Q Q Q S I
GCG AAT TCC TGC AGC CCG GGG GAT CTA CAA CAA CAA TCG ATC ...
EcoRI PstI SmaI
right-hand cloning junction (R)
Y S Q D P S
... TAC TCG cAg GAT cCG TCG ...
BamHI
*

Figure 5.3 Structural features of the GAL4::UNC-119 "bait" plasmid pUBP1, also called pDP#MM143. Genes are not drawn exactly to scale. Below the plasmid are the sequences of the UNC-119 cloning junctions. The actual ligation points are indicated with an asterisk (*), and the destroyed *Bam*HI/*Bgl*II ligations are boxed. Boldface characters indicate the UNC-119-specific sequences. HA, single hemagglutinin epitope; TRP1, tryptophan-1 gene; CYH2, cyclohexamide resistance.

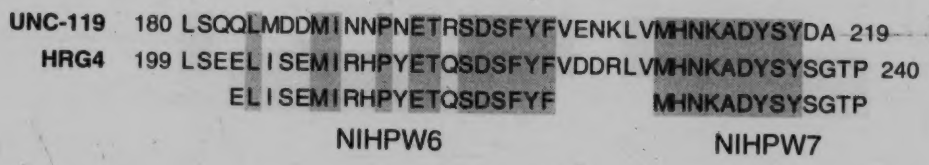

UNC-119 180 LSQQLMDDMINNPNETRSDSFYFVENKLVMHNKADYSYDA 219
HRG4 199 LSEELISEMIRHPYETQSDSFYFVDDRLVMHNKADYSYSGTP 240
ELISEMIRHPYETQSDSFYF
MHNKADYSYSGTP
NIHPW6
NIHPW7

Figure 5.4 Amino acid sequences of the synthetic peptides NIHPW6 and NIHPW7, shown below the carboxyl ends of UNC-119 and HRG4. Conservations between UNC-119, HRG4 and the peptides are shown with darker shading.

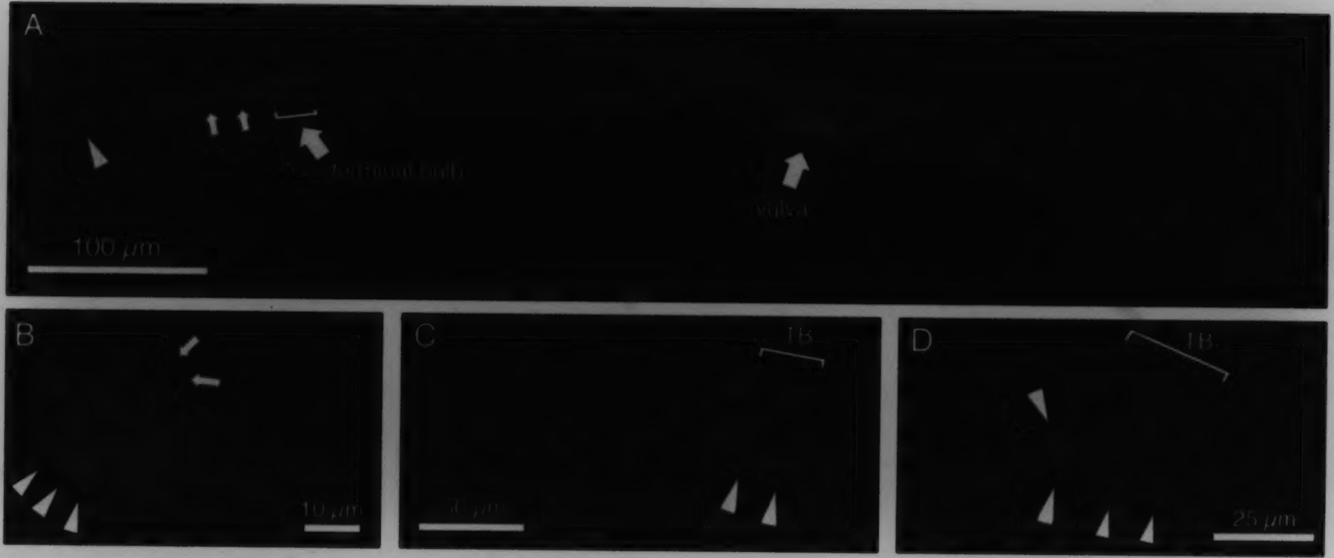


Figure 5.5 Preliminary antibody staining results. **A**, Staining obtained with serum from rabbit immunized with NIHPW6 peptide in strain DP142, *unc-119(ed3)* integrated for the *unc-119*::HRG4 transgene. Note apparently nonspecific staining in vulva and at anterior end (arrowhead). Some staining appears in the terminal bulb of the pharynx, and in a pair of longitudinal structures immediately anterior, which appear to be pharyngeal neurons. Staining with pre-immune serum did not give this pharyngeal staining (data not shown). **B, C** and **D**, *unc-119(e2498)* mutants carrying the HA::UNC-119 fusion plasmid pUHA1 as an extrachromosmal array, stained with anti-HA antibodies. **B**, Late embryo showing staining of sensory structures in the head (arrows) and perinuclear localization in cell bodies of the nerve ring (arrowheads). **C**, Faint nerve ring staining in an adult hermaphrodite. **D**, Perinuclear staining of cell bodies in the nerve ring. Anterior is to the left in all panels. (Continued on next page)

Figure 5.5, continued. Transgenic animals stained with anti-HA antibodies. **E**, *urIs1* strain carrying HA::UNC-6 integrated transgene. Staining is observed in the posterior ventral nerve cord (arrowheads). **F**, Faint structures in vicinity of nerve ring (arrowheads) in same strain as in **B-D**. **G**, Apparent cell bodies in posterior of animal. Anterior is to the left in all panels.

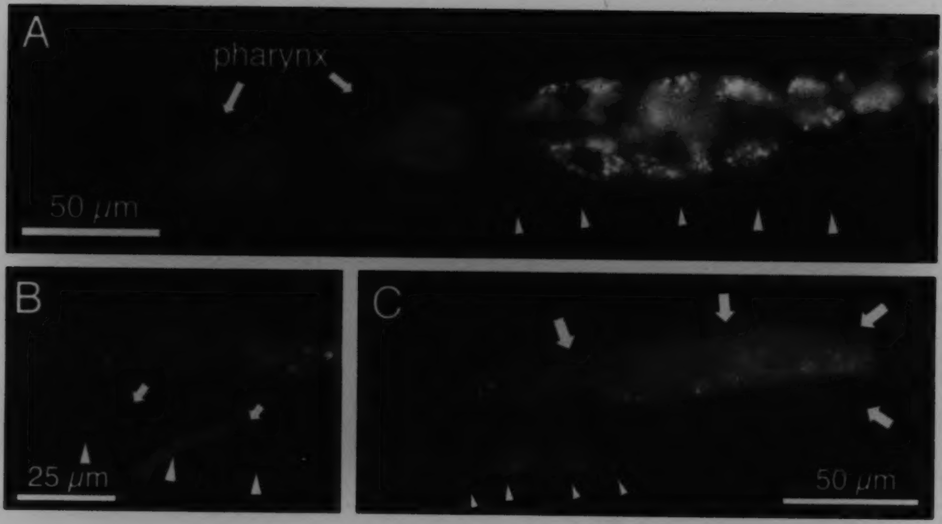


Figure 5.6 Fluorescence resulting from a three-hour 33°C heat shock of wild type animals carrying the pUHS3 [*hsp16-2*::UNC-119::GFP] transgene. Images were made 20 hours after the heat shock treatment. **A**, Note intense punctate accumulation of GFP in the pharynx. There is also diffuse signal in the ventral nerve cord (arrowheads). **B**, Accumulation in the left (arrows) and right (arrowheads) process bundles of the ventral nerve cord. **C**, Apparent expression in the somatic gonad (arrows). The ventral nerve cord near the vulva has also accumulated GFP (arrowheads). Yellow fluorescence visible in panels (A) and (B) is due to gut granules and not GFP accumulation. Anterior is to the left in all figures.

5.5 Bibliography

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. NAR *25*, 3389-3402.

Chalfie, M,. Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. Science *263*, 802-805.

Chien, C. T., Bartel, P. L., Sternglanz, R., and Fields, S. (1991). The two-hybrid system: a method to identify and clone genes for proteins that ineract with a protein of interest. PNAS (USA) *88*, 9578-9582.

Condeelis, J. (1995). Elongation factor 1 α and the regulation of cytoskeletal dynamics, mRNA sorting and growth control. Trends Biochem. Sci. *233*, 169-215.

Cormack, B., Valdivia, R. H., and Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). Gene *173*, 33-38.

Desai, C., Garriga, G., McIntire, S. L., and Horvitz, H. R. (1988). A genetic pathway for the development of the *Caenorhabditis elegans* HSN motor neurons. Nature *336*, 638-646.

Dharmawardhane, S., Demma, M., Yang, F., and Condeelis, J. (1991). Compartmentalization and actin binding properties of ABP-50: the elongation factor-1α of *Dictyostelium.* Cell Motil. Cytoskel. *20*, 279-288.

Fields, S., and Song, O. (1989). A novel genetic system to detect protein:protein interactions. Nature *340*, 245-246.

Harlow, E. and Lane, D. (1988). *Antibodies: A laboratory manual.* Cold Spring Harbor Laboratory Press, NY.

Horvitz, H. R., Brenner, S., Hodgkin, J., and Herman, R. K. (1979). A uniform genetic nomenclature for the nematode *Caenorhabditis elegans*. Mol. Gen. Genet. *175*, 129-133.

James, P., Halladay, J., and Craig, E. A. (1996). Genomic Libraries and a Host Strain Designed for Highly Efficient Two-Hybrid Selection in Yeast. Genetics *144*, 1425-1436.

Jansen, G., Hazendonk, E. Thijssen, K. L., and Plasterk R. H. A. (1997). Reverse genetics by chemical mutagenesis in *Caenorhabditis elegans.* Nature Genetics *17*, 119-121.

Kimble, J., Edgar, L., and Hirsh, D. (1984). Specification of male development in *Caenorhabditis elegans*: The *fem* genes. Dev. Biol. *105*, 234-239.

Kolodziej, P.A. and Young, R.A. (1991). Epitope Tagging and Protein Surveillance. Methods Enzymol. *194*, 508-519.

Link, C. D., Silverman, M. A., Breen, M., Watt, K. E., and Dames, S. A. (1992). Characterization of *Caenorhabditis elegans* Lectin-Binding Mutants. Genetics *131*, 867-881.

Marck, C. (1988). "DNA Strider": a "C" program for the fast analysis of DNA and Protein sequences on the Apple Macintosh family of computers. Nucleic Acids Res. *16*, 1829-1836.

Moerman, D. G., Plurad, S., Waterston, R. H., and Baillie, D. L. (1982). Mutations in the *unc-54* myosin heavy chain gene of *Caenorhabditis elegans* that contractility but not muscle structure. Cell *29*, 773-781.

Murray, J. W., Edmonds, B. T., Liu, G., and Condeelis, J. (1996). Bundling of Actin Filaments by Elongation Factor 1 α Inhibits Polymerization at Filament Ends. J. Cell Biol. *135*, 1309-1321.

Ogura, K., Shirakawa, M., Barnes, T. M., Hekimi, S., and Ohshima, Y. (1997). The UNC-14 protein required for axonal elongation and guidance in *Caenorhabditis elegans* interacts with the serine/threonine kinase UNC-51. Genes Dev. *11*, 1801-1811.

Stringham, E. G., D. K. Dixon, D. Jones, and E. P. M. Candido. (1992). Temporal and spatial expression patterns of the small heat shock (*hsp16*) genes in transgenic *Caenorhabditis elegans*. Mol. Biol. Cell *3*, 221-233.

Wadsworth, W. G., Bhatt, H., and Hedgecock, E. M. (1996). Neuroglia and Pioneer Neurons Express UNC-6 to Provide Global and Local Netrin Cues for Guiding Migrations in C. elegans. Neuron *16*, 35-46.

Williamson, T., Gordon-Weeks, P. R., Schachner, M., and Taylor, J. (1996). Microtubule reorganization is obligatory for growth cone turning. PNAS (USA) *93*, 15221-15226.

Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., *et al.* (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome *III* of *C. elegans*. Nature *368*, 32-38.

6 General Discussion

6.1 Introduction

The individual chapters of this thesis contain discussions pertaining directly to the experimental approaches used and their conclusions. As the text from Chapters 2 through 4 have been published over a several year period, the discussions fall short of putting all *unc-119* data into a global perspective, and are in some respects outdated (particularly Chapter 2). This chapter attempts to describe the sum total of the thesis, draw conclusions that emerge from the work as a whole, and make predictions about biological roles for UNC-119 and its homologues that were too speculative to be included in published papers. Furthermore, as the results also raise important questions, I will describe experimental strategies.

6.2 The basis for the *unc-119* mutant phenotype

6.2.1 unc-119 mutations

All known *unc-119* alleles were found to be recessive, unconditional amorphs, based on the similar phenotypes of *unc-119/unc-119* and *unc-119/tDf2* combinations, and the wild-type phenotype of *unc-119/+* animals (Chapter 2). The nature of the molecular lesions in each allele was consistent with inactivation of UNC-119 function (Figure 2.16). This facilitated analysis of the *unc-119* mutant phenotype, because the defects were uniform among all alleles and attributable to complete loss of *unc-119*. With a more detailed understanding of the phenotype at the cellular level (Chapter 4) we can propose explanations for the behavioural defects in *unc-119* mutants.

6.2.2 The basis for loss of coordinated movement

The most obvious phenotype of the different *unc-119* alleles was the loss of wild-type movement. Younger larvae could move somewhat better than adults, which exhibited a paralyzed ventral coiling shape (Chapter 2). The wild-type muscle ultrastructure and post-synaptic response to levamisole were consistent with a defect in the nervous system. Examination of neuron structure using Green Fluorescent Protein (GFP) reporters and anti-GABA antibodies demonstrated multiple structural defects in axons. These axons, which must allow growth cone migration in the absence of axon-axon fasciculation cues during development (McIntire *et al.*, 1992), appeared to be the most affected in *unc-119* animals. With these data, we can explain the loss of coordinated locomotion in *unc-119* mutants.

Commissural axons are a component of four of the six classes of neuron mediating coordinated locomotion in *C. elegans* (Figure 1.10; White *et al.*, 1986; White *et al.*, 1992). We might therefore expect *unc-119* mutants to have abnormal DB, VD, DD and DA motor neurons, but relatively normal VB and VA motor neurons, although this was not experimentally confirmed. Neurons of the VB and VA classes are excitatory, and make neuromuscular junctions in the ventral nerve cord with ventral body wall muscle arms (Driscoll and Kaplan, 1997). Therefore, only the ventral muscle quadrants can receive signals in an *unc-119* mutant, and these signals would only be stimulatory, not inhibitory. This would be expected to only allow ventral body wall muscle contraction, which is entirely consistent with the ventral coiling seen in *unc-119* animals (Figure 6.1).

There are two potential compensatory mechanisms that might cause phenotypic variability. First, despite aberrant outgrowth, some commissural axons may still extend to the dorsal nerve cord. Second, muscle arms can find misplaced motor neurons, which may allow them to receive synaptic input (Figure 1.13; reviewed by Jorgensen and Rankin, 1997). However, the relatively uniform degree of the locomotory defect in *unc-119* mutants (e.g., Figure 2.2) suggests that these mechanisms only contribute to a small degree of variability.

6.2.3 The dauer formation and amphid dye-filling defects

Mutations in *unc-119* result in animals that cannot form dauer larvae under starvation conditions (Chapter 2). Epistasis analysis allowed the placement of *unc-119* in the dauer formation pathway, at a step corresponding to mutations affecting dye filling (Dyf) of the amphid neurons (Figure 2.13). The results were consistent with the dye-filling defect observed in *unc-119* animals (Chapter 4), as such mutants are unable to form dauer larvae (Starich *et al.*, 1995). All Dyf animals examined by electron microscopy have defective amphid sensory cilium structure (Perkins *et al.*, 1986). By extension, *unc-119* animals are likely to have a similar structural defect in the amphids, consistent with the abnormal neuron morphology seen for body wall neurons (Chapter 4).

6.2.4 Basis for the egg laying and constitutive pharyngeal pumping defects

The weak egg laying defect observed in *unc-119* mutants suggested a defect in the egg laying circuitry (Chapter 2). The finding that serotonin, but not imipramine, could potentiate egg laying in hermaphrodites was consistent with a presynaptic defect (Chapter 4). Furthermore, anti-serotonin antibodies still stained the HSN cell bodies, which were found to be at the wild-type position in *unc-119(e2498)* animals (Figure 4.17). This would argue that the innervation of the egg laying muscles was disrupted, perhaps by structural defect in the HSNs. However, a small but significant 'constitutive' level of egg laying persisted, suggesting that mutants still had some capacity to stimulate the egg laying muscles. This is not necessarily a contradiction; VC neurons also synapse onto the egg laying muscles, but their role in egg laying, though unessential, is unknown (White *et al.*, 1986; Desai *et al.*, 1988).

Two unrelated genes, *unc-34* and *unc-73*, also show weak constitutive egg laying that has been proposed to result from abnormal negative regulation of the HSNs (Desai *et al.*, 1988). It may be that abnormal negative regulation of egg laying occurs in *unc-119* animals. The observation that *unc-119* mutants also exhibit constitutive pharyngeal pumping is consistent with this hypothesis. Stimulation of pharyngeal pumping and egg laying, and depression of locomotion, all occur in the presence of food (Croll, 1975; Croll and Smith, 1978). Addition of exogenous serotonin causes a similar response (Horvitz *et al.*, 1982). The NSMs (neurosecretory motor neurons) are a candidate pair of neurons that mediate this response to food by releasing serotonin into the pseudocoelom: the NSMs stain with anti-serotonin antibodies, and have putative sensory endings in the pharynx, where bacteria accumulate (Albertson and Thomson, 1976).

As another set of chemosensory neurons, the amphids, are suspected to have a structural defect in *unc-119*, an analogous defect may exist with the NSMs. This might result in constitutive stimulation of behaviours that occur in the presence of food, namely pharyngeal pumping and egg laying. The repression of dauer formation is also consistent with a constitutive false perception of food.

The complex interactions among neurons mediating these behaviours preclude an unequivocal assignment of a specific defect as the basis of the egg laying phenotype. Indeed, the widespread neural expression of *unc-119* implies that we may expect different aspects of egg laying to be perturbed. Visualization of the anterior HSN processes in *unc-119* mutants may provide a partial answer; abnormal anterior axons would result in misregulation of the HSNs (Desai *et al.*, 1988).

6.3 Conclusions about *unc-119* gene expression

6.3.1 unc-119 reporter transgenes

The expression pattern of an *unc-119::lacZ* transgene, described in Chapter 2, suggested that the *unc-119* gene is expressed in the majority of nervous system cells. This was supported by observations using reporter fusions to GFP (Chapter 4), and by similar transgenes constructed from the *C. briggsae unc-119* promoter (Chapter 3). The reporter fusions also contained 5' flanking sequences found in genomic clones that could completely rescue the *unc-119* defects. The reproducibility of the expression patterns, especially using promoters derived from two *Caenorhabditis* species, suggests they are not artifactual. As well, the use of a larger promoter gave similar expression (§4.14).

The expression of a transgene, however, can only be taken as indirect evidence for the native expression of a gene, due to the nature of *C. elegans* transgenes, which form multicopy arrays (Mello *et al.*, 1991; Mello and Fire, 1997). Determination of the precise pattern of *unc-119* expression will have to rely on *in situ* analysis, or detection of UNC-119 protein. Although the best way to confirm requirement for UNC-119 within neurons would involve mosaic analysis (Herman, 1997), some confidence can be placed in the spatial pattern of *unc-119* reporter transgenes by correlation with the mutant phenotype (§6.6).

The temporal expression of *unc-119* transgenes suggests that the *unc-119* mRNA accumulates in the early embryo and is expressed throughout adulthood. The identity of the first *unc-119*::lacZ expressing cells was not determined, however. It should be possible, in an integrated *unc-119*::GFP strain, to identify the precise cells expressing the reporter by recording the growth of embryos. The pattern of expression in older embryos predicts that these cells will be neural cell precursors.

6.3.2 unc-119 transcripts

The data support the existence of a single *unc-119* gene product. No coding potential was found elsewhere in the pDP#MM016 rescuing clone (Chapter 2), and only one open reading frame exists in the overlapping cDNAs recovered. A similar situation was found with the *C. briggsae* gene (Chapter 3). Preliminary Northern blot analysis of mixed-stage total RNA preparations in both *C. elegans* and *C. briggsae* were consistent with the production of a single message (Chapter 3), and RT-PCR performed in both species only generated gene-specifc products predicted by amplification of a single cDNA (Chapters 2 and 5).

Even if alternative transcripts exist, only a single message appears to be necessary. A 'mini-gene' fusion of the *unc-119* promoter to the *unc-119* cDNA (pDP#MM051) was capable of rescuing the mutant defects in *eDf2; eDp6 III*, a strain deleted for *unc-119* (Chapter 2). As further indirect evidence for requirement of a single polypeptide, the single HRG4 and DmUNC119 coding regions were able to rescue *unc-119* null mutants (Chapter 4).

6.4 The set of *unc-119* alleles

All of the extant alleles of *unc-119* result in apparent loss of function, an observation that is borne out by the nature of the molecular lesions. The canonical allele *e2498* contains a transposon insertion in the coding region; two alleles, *ed3* and *ed4*, are nonsense mutations occurring within ten codons of each other; the mutation *ed9* results in a frameshift; and the chromosomal rearragements *eDf2* and *eDp6 III* delete the entire coding region (Figure 2.16).

While the collection of null alleles has been convenient in allowing an unequivocal description of the loss-of-function phenotype, a temperature sensitive (ts) conditional allele would be useful for the study of temporal gene requirement and, potentially, the identification of putative interacting proteins through the generation of extragenic suppressors. For example, rare mutations in *unc-54* can suppress mutations in *unc-22* (Moerman *et al.*, 1982).

Attempts were made to generate conditional alleles for the *unc-119* that relied on screening the F_1 progeny of a noncomplementation screen at 25°C (Chapter 2). These screens met with failure (although one produced the unconditional allele *ed9*) perhaps owing to several factors. The mutagen EMS, which predominantly generates GC to AT transitions, was used in each case; as most temperature-sensitive alleles of genes are of the missense type, the use of this mutagen constrained the type of amino acid substitutions that could be generated. Furthermore, the *unc-119* gene appears to be a smaller-than-average target for mutagenesis, which may explain why it was not identified by other investigators (Chapter 2). However, this might simply be because EMS-induced ts mutations are impossible (or extremely rare) in *unc-119*. This mutagenic bias could be alleviated with the use of *N*-ethyl-*N*-nitrosourea (ENU), which has been found to cause a greater spectrum of mutations in *C. elegans* (De Stasio *et al.*, 1997).

An alternative to the creation of an endogenous conditional allele is to introduce a conditional transgene into a mutant background. The transgene may be a synthetic temperature-sensitive (ts) allele, anticipated to have this property based on the existence of a 'homologous' ts mutation in another system; however, there are no known mutations in any *unc-119* homologs.

We can speculate as to what type of mutations would be likely to result in temperature-sensitive function of UNC-119. The molecular lesions of the extant *unc-119* alleles are consistent with the finding that the human homologue HRG4, which can rescue the *unc-119* mutant phenotype, requires the very carboxy terminus for function. A synthetic truncation immediately prior to the protein segment -VHMNHKADY-, which is conserved in *C. elegans* UNC-119, completely abolished the ability of HRG4 to rescue *unc-119* mutants (Chapter 4). However, deletion of the unconserved amino termini did not interfere with rescuing ability of HRG4 or DmUNC119, suggesting that the corresponding amino-terminal segment of UNC-119 is not vital. We may extend this to predict that the spacer region between segments A and B, and any amino acid positions that have unconserved changes between species, are not required *per se* for UNC-119 function. The sum total of all such ‘flexible’ positions in UNC-119 define missense mutation targets that would probably go undetected in a mutagenic screen, and that are therefore unlikely to result in a conditional phenotype.

The foreknowledge that the carboxy terminus is absolutely required for HRG4 function, and by extension, UNC-119 function, enables a directed method by which to generate potential conditional alleles. Oligonucleotide primer pools, encoding amino acid substitutions, could be generated that overlap the carboxyl portion of the open reading frame (ORF), and provide a restriction enzyme site for the cloning of modified coding regions obtained by polymerase chain reaction (PCR). A plasmid backbone (pDP#MM016MM), used to test function of the human and *Drosophila* cDNAs in *C. elegans*, is already available. Such a conditional transgene, if found, could be integrated and used as a virtual “gene replacement.” There are two potential problems with this approach. First, a weaker form of UNC-119 might still fully function as a transgene due to compensation by overexpression in a multicopy array (Mello, 1991). Second, the conserved residues might be absolutely required for function.

6.5 The partial rescue of *unc-119* by an incomplete *unc-119* gene

The observation that an incomplete fragment of the *unc-119* gene (in plasmid pDP#MM019) can contribute partial rescue of the *unc-119* movement defect deserves special mention. It is the only case of such partial rescue ever obtained for *unc-119* that cannot be attributed to transgene mosaicism. The *unc-119* fragment in pDP#MM019 lacks the *unc-119* promoter and first two exons; its left-hand end starts at an *Eco*RI site in the middle of the second intron. The fact that partial rescue is obtained at all suggests that expression of the remaining portion of *unc-119* must be occurring, perhaps from cryptic transcriptional initiation sites within the intronic sequences preceding exon III. It could be assumed that the last two introns, whose donor and recipient splice sequences are intact, are processed normally; the start of exon III contains an in-frame ATG start codon that might be used for translational initiation of an amino-truncated form of UNC-119 (Chapter 2).

A contradiction emerges in light of results with the *unc-119*::HRG4 and *unc-119*::DmUNC119 transgenes. These experiments demonstrated the ability of two distant UNC-119 homologues to fully rescue *unc-119* in the absence of a large portion of their amino termini. Only the amino terminal 20 amino acids are excluded in pDP#MM019, and the remaining portion of UNC-119 retains even more than the 'minimal' portion defined by rescue/conservation with HRG4 and DmUNC119.

There are three possible reasons. First, the *C. elegans* gene might require the full amino terminus. This is certainly consistent with the 90% identity between the “disposable” amino termini of *C. elegans* and *C. briggsae*. Second, translational initiation of the foreshortened UNC-119 may be weak, or occur at a point farther into the coding region, excluding conserved (and likely functionally essential) sequences from the polypeptide. Third, cryptic transcriptional initiation of the presumed truncated message might be too weak to allow the transgene to fully rescue.

The first alternative does not seem likely given the full rescue by an amino-truncated HRG4 cDNA*. The second and third alternatives could be distinguished by expressing an amino-truncated UNC-119 under control of the full *unc-119* promoter. It seems most likely, however, that weak expression, either by poor transcriptional or translational initiation, accounts for failure of pDP#MM019 to fully rescue the *unc-119* movement defect: the intron fragment is not likely to contain good neural *cis*-acting initiation sequences. An obvious way to resolve this question is to add the normal *unc-119* promoter (i.e. from vector pDP#MM065, §4.13) to the *unc-119* fragment in pDP#MM019. The results with HRG4 and DmUNC119 suggest that such a clone should enable complete rescue in an *unc-119* mutant.

6.6 Correlation of *unc-119* expression and the mutant phenotype

Superficially, the neural expression of *unc-119* reporter transgenes is consistent with the neural phenotype of *unc-119* mutants (Chapters 2,3,4): defective axons are revealed in *unc-119* mutant animals expressing an *unc-119*::GFP fusion. Expression in the ventral nerve cord is consistent with defective process fasciculation (Chapters 2,4); expression in commissural axons is consistent with their abnormal outgrowth (Chapter 4); and expression in the amphids is consistent with the dye-filling defect (Chapters 2,4). However, there are some cases of transgene expression that are not fully understood, either because the expressing cells have not been identified, or because there is no detectable mutant phenotype involving them.

For example, *unc-119*::GFP and *unc-119*::lacZ expression were seen in large, leaf-like structures in the head (Chapter 2, Chapter 4). The identity of these cells was not determined, although they appeared to be associated with the amphids, and may therefore be socket or sheath cells (White *et al.*, 1986). The *unc-119*::lacZ reporter, which has a nuclear localization signal, also stained some additional unidentified nuclei in the head (Chapter 2) that could belong to these cells.

The *unc-119*::GFP reporter was also seen in pharyngeal neurons. There was no apparent defect in pharyngeal pumping (other than its regulation), which suggests that the requirement for *unc-119* is not as stringent in these neurons, or that the expression is artifactual. Laser ablation of pharyngeal neurons has shown that of the 20 neurons in the pharynx, only three (M4, MC and M3) are required for nearly normal feeding (Avery and Thomas, 1997). Therefore, defects in the pharyngeal neurons in *unc-119* mutants may not affect pumping ability. Alternatively, there may be a requirement for UNC-119 within pharyngeal neurons to allow the neurons to respond to inhibitory signals.

A similar problem arises in extrachromosomal arrays of some UNC-119::GFP fusions, which gave punctate staining associated with muscle dense bodies in the anterior of the animal (Chapter 4). There were no apparent muscle defects in *unc-119* mutants, although muscles in the very anterior of the animal were not been examined (Chapter 2). There is a precedent for neural genes giving similar punctate muscle expression. Antibody detection of VAB-8 reveals dense-body associated staining throughout the animal, without a concomitant muscle defect in *vab-8* mutants (G. Garriga, pers. comm.), and antibodies to the

EGL-10 protein stain neural processes and punctate regions in body wall muscle in wild-type animals (Koelle and Horvitz, 1996). As other *unc-119* transgenes do not cause punctate staining, immunological detection of UNC-119 will be required to settle the question of whether or not the endogenous protein is found in (or associated with) muscle dense bodies.

Lastly, accumulation of GFP in hermaphrodite distal tip cells (DTCs) was observed in *unc-119*::GFP animals that had been subjected to fixation and collagenase treatment (Chapter 4). It is strange that this expression had not been noticed before; presumably, the disruption of the cuticle and spreading out of the gonad permitted these cells to be seen.

The DTCs guide the growth of the gonad during hermaphrodite and male development, and are involved in the subsequent production of gametes (Kimble and White, 1981). During gonad developmen, the DTCs appear to respond to gradients provided by netrin/*unc-6* cues: the DTCs express *unc-5* when they migrate dorsally (M. Su *et al.*, in prep., cited in Antebi *et al.*, 1997). However, no gonadal defects were seen in *unc-119* mutants, although there was a general decline in brood size (Chapter 2). Therefore, if *unc-119* is expressed in the DTCs, it may be that its function overlaps that of another gene.

6.7 The functional and structural conservation of UNC-119

Perhaps the most gratifying finding of this work was that *unc-119* has homologues in vertebrates and in the fruit fly *Drosophila* that are capable of functioning in place of the *C. elegans* gene. The extent of conservation was quite high, suggesting a common ancestral origin for these genes. Consistent with this, an intron was found at the same site in the coding regions of HRG4, DmUNC119 and the *Caenorhabditis unc-119* genes (Figure 4.11).

If the UNC-119 protein has a general function in the nervous system, rather than in photoreception (which seems to be the case), the apparent retina-specific expression of the human and *R. norvegicus* homologues of *unc-119* (Higashide *et al.*, 1996) is somewhat puzzling. It is possible that there are no UNC-119-like proteins in vertebrate brain and spinal cord as a result of the higher complexity of these neurons; however, the nematode UNC-6 family is structurally conserved with the vertebrate Netrin family, which are expressed in the developing spinal cord (Ishii *et al.*, 1992; Kennedy *et al.*, 1994). More than likely, then, there are other vertebrate UNC-119 homologues to be found. Indeed, within *C. elegans*, a related gene is predicted to exist, C27H5.1 (Chapters 3 and 4).

The sequence conservation and the ability of divergent homologues to rescue suggest that proteins of the UNC-119 class interact with other well-conserved proteins. These factors could be ubiquitous, such as actin or microtubules, or they might be conserved proteins found only in neural cell types, such as UNC-33/CRMP-62 (Goshima *et al.*, 1995). This predicts that there are undiscovered *C. elegans* genes with vertebrate homologues expressed (at least) in the retina, and may reveal as-yet unknown shared mechanisms of neurogenesis.

Indeed, there are vertebrate retina homologs of other neural *C. elegans* genes. The *tax-2* and *tax-4* gene products are structurally similar to components from cyclic nucleotide-gated channels from olfactory neurons and vertebrate photoreceptors (Coburn and Bargmann, 1996; Komatsu et al., 1996), and consistent with their predicted role, are expressed only in sensory neurons of *C. elegans* (Coburn and Bargmann, 1996). In addition, expression of *ceh-10*, a homolog of the vertebrate retina-specific homeobox genes Chx10 and Vsx-1, has been observed in a specific set of nematode neurons (Svendsen and McGhee, 1995).

Finally, the predicted gene C27H5.1 (Chapter 3) is structurally related to a vertebrate phosphodiesterase subunit, whose expression is apparently retina-enriched like that of HRG4 (D. Pilgrim, pers. comm.; Florio *et al.*, 1996). The expression of a C27H5.1::GFP fusion (§4.4.2) is consistent with a neural role in *C. elegans.* Although there is only limited similarity between UNC-119 and C27H5.1 (Figure 3.5), a common role in regulation of signal transduction may be suggested for the extended UNC-119/C27H5.1 family of proteins.

6.8 A possible role for UNC-119 in growth cone extension

Cues that provide growth information during commissural outgrowth are largely different from those that act in fasciculation. Mutations have been isolated in *C. elegans* that affect either fasciculation or commissural outgrowth but not both (McIntire *et al.*, 1992). This is explained by the different mechanisms occurring in each type of axonogenesis: Fasciculation requires perception of local cues, largely determined by cell-cell adhesion, while commissural outgrowth also relies on reception of long-range cues (Figure 1.14; Tessier-Lavigne and Goodman, 1996).

The presence of both commissure and fasciculation defects in *unc-119* mutants suggests that a cellular function common to both phenomena is affected. Specifically, the intracellular response to growth signals, which occurs within axons during both fasciculation and commissural outgrowth (Tessier-Lavigne and Goodman, 1996), seems a likely place at which UNC-119 could be acting. If this is true, UNC-119 should be found within neurons, either in growth cones or in developing axons.

There are several lines of evidence that indirectly support a cytoplasmic localization for UNC-119. First, *unc-119* reporter fusions gave expression in the same neurons that showed structural defects in mutants, suggesting that UNC-119 functions in the same cells as it is expressed, and that it is not required for neural differentiation. Second, functional GFP fusions of both UNC-119 and the human homologue HRG4 can target GFP to axons. Third, the functions of UNC-33/CRMP-62 and UNC-51, two proteins within the same set of genes that share *unc-119*-like defects (McIntire *et al.*, 1992; Chapter 4) are known to be cytoplasmic (Goshima *et al.*, 1995; Ogura *et al.*, 1994). Finally, the lack of any obvious signalling motifs in the sequences of UNC-119, DmUNC119 and HRG4 is consistent with cytoplasmic localization *in vivo* (though not conclusive of one).

Localization of UNC-119 to axons, and a mutant defect in outgrowth and fasciculation, suggest that this protein functions intracellularly to allow neurites to develop properly. Growth cones, the specialized termini of developing neurons, migrate to extend axons through a three-stage cycle of exploration, site selection, and stabilization/axon formation (§1.10; reviewed in Tanaka and Sabry, 1995). Cytoplasmic projections (filopodia and lamellipodia) 'sense' the growth cone microenvironment and allow the growth cone to extend the axon in the proper direction. The many different steps and proteins involved in growth cone elongation predicts a diversity in the particular types of defects likely to be seen as a result of mutation. For example, foreshortened or misrouted neurites are seen in *unc-6*, *unc-5* and *unc-40* mutants, which affect the presence or reception of positional information required for guidance (Chan *et al.*, 1996; Ishii *et al.*, 1992; Leung-Hagesteijn *et al.*, 1992; Wadsworth and Hedgecock, 1996). In contrast, *unc-33* mutants display highly branched commissures (Figure 4.4), and consistent with a cytoskeletal abnormality, show excess numbers of microtubules in sensory neurons (Li *et al.*, 1992; Hedgecock *et al.*, 1985). The similarity of other neural defects in *unc-33* with those of *unc-119*, described in §4.4.1, suggests that these mutants may share an underlying defect.

How could UNC-119 be functioning in axon outgrowth? One possibility is that the branching of *unc-119* commissures results from a defect in the ability to collapse inappropriate filopodial extensions, or to correctly favor a selected extension (Figure 6.2). As site stabilization requires reorganization of microtubules (Tanaka and Sabry, 1995; Williamson et al., 1996), UNC-119 may be a factor that influences microtubule assembly or disassembly. As such, it could provide a link between recognition of external guidance factors and internal cytoskeletal rearrangements accompanying axonogenesis (Tanaka and Sabry, 1995; Tessier-Lavigne and Goodman, 1996). Confirming such a model will likely need new biochemical or genetic data to identify proteins interacting with UNC-119. As an alterative, observation of migrating growth cones in *unc-119* animals, with a suitable marker, might reveal mechanistic differences in axonogenesis.

The apparent expression of *unc-119* in the distal tip cells (DTCs) is somewhat consistent with a cytoskeletal function, if we assume that expression occurs during DTC migration. As cell migration and axon outgrowth both involve rearrangement of the cytoskeleton in response to external cues, this could be another example of how UNC-119 functions to establish changes in cell shape.

A cellular mechanism by which UNC-119 may influrence the cytoskeleton was suggested by preliminary results from a yeast two-hybrid screen (Chapter 5). The recovery of two overlapping cDNAs for Elongation Factor 1 α (EF-1α) suggested this protein might interact with UNC-119 *in vivo*, although the conclusive experiments have not been performed. The conserved protein EF-1α has been shown to affect polymerization of actin both positively and negatively, depending on its relative molar concentration (Murray *et al.*, 1996). Microtubule polymerization and depolymerization are required for proper growth cone function (Williamson *et al.*, 1996), and actin polymerization is important for the extension of lamellipodia and filopodia (reviewed in Tanaka and Sabry, 1995). Therefore, a role for UNC-119 in the modulation of actin and microtubule stability, via EF-1α, is an intriguing possibility.

Finally, it may be possible to rule out a direct role for UNC-119 in the regulation of G protein signalling in neurons. None of the *unc-119* transgenes showed any obvious phenotypic consequences, despite the known overexpression of genes in extrachromosomal arrays (Mello, 1990). In contrast, at least two genes that regulate neural signal transduction show dosage effects. Reporter GFP fusions of the gene *goa-1*, which encodes an α subunit of G_o, show expression in muscles and neurons associated with serotonin-mediated behaviour (Ségalat *et al.*, 1995). Behaviours such as locomotion and egg laying are stimulated in *goa-1* loss of function mutants, but inhibited in animals carrying an extrachromosomal *goa-1* transgene (Ségalat *et al.*, 1995). The opposite is true for *egl-10*: mutants show decreased egg laying and locomotion, while animals carrying an *egl-10* transgene display an increase in these behaviours (Koelle and Horvitz, 1996). The EGL-10 protein is found in neural processes, and has been proposed to regulate G protein signalling (Koelle and Horvitz, 1996). Therefore, the lack of a behavioural transgene effect in *unc-119* suggests it does not have a regulatory role similar to *egl-10* and *goa-1*.

In summary, the specific cellular phenotype of *unc-119* mutants (Chapter 4), the neural expression of reporter transgenes (Chapters 2-4), and the cytoplasmic localization of functional GFP fusions within neurons (Chapter 4), suggest a role for UNC-119 in the establishment of the neural cytoskeleton. This role may be structural, in that UNC-119 may be part of the cytoskeleton itself. Alternatively, UNC-119 may have a regulatory role, either in the transduction of guidance signals received at the membrane, or in the manifestation of cytoskeletal changes associated with those signals.

As a supplementary note to the putative role of UNC-119, *in situ* analysis of expression of the *Drosophila* homologue DmUNC119 has revealed expression in the central nervous system in embryos (R. Jacobs, pers. comm.). This suggests that UNC-119 and DmUNC119 may share similar functions in the developing nervous system.

6.9 A role for UNC-119/HRG4 in maintenance

The early onset of expression of RRG4 in the rat retina (Higashide *et al.*, 1996), and of *unc-119*::GFP in the *C. elegans* embryo, long before functional photoreceptors and neurons are formed, is consistent with a role for these proteins in development. However, RRG4 mRNA persists in adult retina (Higashide *et al.*, 1996). In *C. elegans*, adults express *unc-119*::GFP and *unc-119*::lacZ fusions (even under control of the *C. briggsae* promoter; Chapters 2 and 3). With the caveat that expression of a transgene fusion need not imply function or requirement, this similarity in temporal expression does suggest that UNC-119/HRG4 proteins also play a role in maintenance or function.

This role need not be different from that occurring during development. It is not unreasonable to propose that the cytoskeleton remains somewhat dynamic, even in fully formed *C. elegans* axons; in vertebrate neural development, axons undergo extensive 'optimization' even after they have made synapses with their targets (reviewed in Nicholls *et al.*, 1992). Therefore, there may be a continuing role for UNC-119 in maintaining the cytoskeleton.

A proposed maintenance role for UNC-119 predicts that loss of *unc-119* activity in adult worms would cause some type of progressive neural function defect, as nervous system development would already be complete. This experiment has not been done, as all extant alleles of *unc-119* are unconditional nulls (Chapter 2). Furthermore, as of the writing of this thesis, no known mutations exist for any of the UNC-119 homologues.

6.10 Future work on *unc-119*

Our laboratory is continuing research into the role of UNC-119 and its homologues. First, the preliminary results from a yeast two-hybrid screen suggest that it will be possible to identify proteins that interact with UNC-119. The availability of strains rescued with HA-tagged versions of UNC-119 should expedite co-immunoprecipitation studies from intact worms to confirm putative interactions. A putative role for UNC-119 in cytoskeletal rearrangement suggests that there may be interactions with other proteins with similar defects, such as UNC-33 or UNC-51; a directed two-hybrid screen of such cloned genes may fortuitously identify an interacting protein.

Second, the construction of several GST fusions will allow expression of fusions in *E. coli* for purification and the generation of polyclonal antisera. These can be used to study the *in vivo* localization of the UNC-119 polypeptide, and will afford another means by which to assess protein-protein interactions. Specifically, it will be of interest to determine if there is punctate localization of UNC-119 in axons.

Third, it should be possible to generate more alleles of *unc-119* using the screen described in §2.7. Specifically, temperature-sensitive alleles will be of interest as these will be useful in establishing the temporal requirement for *unc-119* function. Alternative mutagens, such as ENU, might be tried to eliminate the mutational bias that appeared to result from use of EMS.

Fourth, the existence of a homologue within *C. elegans* suggests UNC-119 may be part of a larger extended family of neural proteins. Deletions can be generated in the C27H5.1 predicted gene to determine its possible role, if any, in *C. elegans* nervous system development or function. Combination of this mutation with an *unc-119* null mutation may reveal overlapping function in the two genes.

Fifth, the identification of a *Drosophila* homologue may allow the first analysis of a mutation in an UNC-119 gene outside of *C. elegans*. There is reason to believe that such a mutation would affect axon outgrowth in *Drosophila*: a DmUNC119 cDNA can replace *unc-119* function in *C. elegans*, and this gene appears to be expressed in the *Drosophila* nervous system (R. Jacobs, pers. comm.). A vertebrate mutation, such as a mouse knockout, or the correlation of mutation in HRG4 with a human disease, will also serve a similar purpose. As the phenotype in *C. elegans* is understood at a cellular level, it will be of interest to see if mutation in other systems causes similar defects.

Sixth, the identification of the corresponding breakpoint sequences in *eDp6* may shed light on mechanisms of chromosome repair in *C. elegans*.

Finally, the degenerate UNC-119/HRG4 oligonucleotides have already been useful for the identification of homologues from diverse phyla. Approximately one-half of the species whose DNA has been subjected to PCR using UNC-119 degenerate primers have yielded an UNC-119 homologue. To date, species in which degenerate PCR products have been sequenced and verified to contain an UNC-119-like partial ORF include zebrafish (A. Manning, pers. comm.), the related nematode species *C. remanei* , and a primitive spinculate worm (D. Pilgrim, pers. comm.). If the UNC-119 class of proteins is universally conserved in animals, construction of alternate primer sets may enable the identification of more UNC-119 homologues.

6.11 Concluding remarks

The results of experiments presented in this thesis allow us to conclude that:

1. The *unc-119* gene of *C. elegans* contributes to neural development. Mutants display abnormal axon outgrowth and fasciculation concomitant with defects in many behaviours. Muscle appears to be unaffected in *unc-119* animals.
2. Loss of function mutations in *unc-119* can be generated using the mutagen EMS. Three new mutants, *ed3*, *ed4* and *ed9* were generated in straightforward screens.
3. The *unc-119* gene has been cloned from both *C. elegans* and *C. briggsae*. The genes each express a message of approximately 1 kb in size and encode proteins of size 219 aa and 217 aa, respectively. The *C. elegans unc-119* EMS mutants were found to have lesions resulting in trucated open reading frames. The original allele, *e2498*, carries a Tc1 insertion in a coding exon.
4. The UNC-119 protein is novel, and functionally conserved with three structurally related proteins: the UNC-119 homologue from the nematode *C. briggsae*, the human homologue HRG4, and the newly identified homologue in *Drosophila*, DmUNC119. This suggests that UNC-119 and its homologues constitute members of a large group of related neural proteins.
5. The deletion *eDf2* is a large interstitial deletion that removes approximately 1/3 of the DNA on the right of LG *III*, including most of the *unc-119* gene, but retains the very right end of the chromosome. A complex model for the concomitant generation of *eDf2* and *eDp6* is implied.

6. The biochemical function of UNC-119 and its homologues is unknown. The cellular defect suggests a role in the cytoskeleton or in regulating cytoskeletal rearrangements in response to neural growth cues.
7. Plasmids generated for future protein work (i.e. yeast two-hybrid bait plasmid and GST/HA fusion plasmids) may identify UNC-119 interactions, and allow the generation of polyclonal antibodies to study *in vivo* localization.
8. The study of the structure, expression and function of UNC-119 homologues, particularly in *Drosophila*, may reveal other roles for this class of proteins.

The surprising turns taken by the experiments performed during the tenure of this thesis have underscored the excitement that basic research can elicit in model systems. The work on UNC-119 has raised interesting questions about neurogenesis, not only in *C. elegans*, but in *Drosophila* and vertebrates. It is my sincere hope that this work will be followed up, and that it will continue to yield exciting insights into fundamental mechanisms of nervous system development.

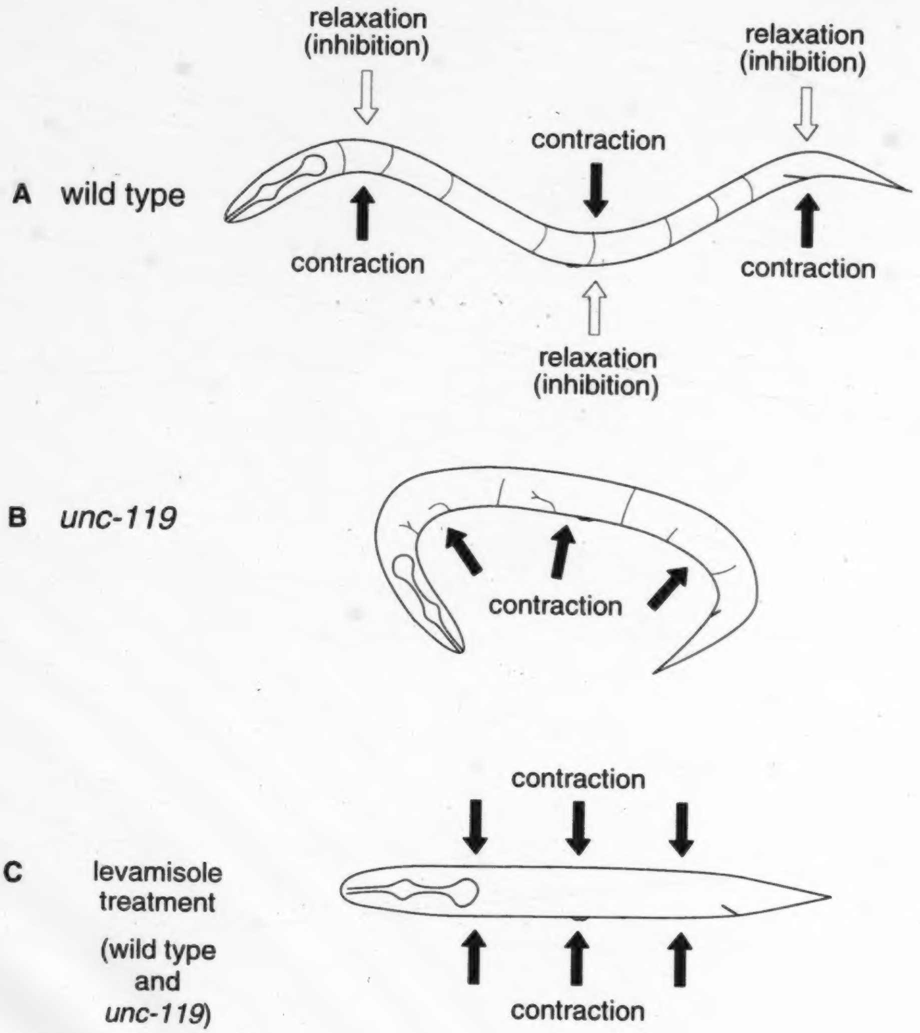


Figure 6.1 Proposed basis for the uncoordinated locomotion of *unc-119* mutants. **A**, In the wild type, sinusoidal locomotion results from longitudinally propagated waves of body wall muscle contraction. **B**, In *unc-119* animals, four of the six classes of motor neuron are defective due to improper outgrowth of commissural axons. The remaining VB and VA motor neurons only permit contraction of the ventral muscles, which results in ventral coiling. **C**, Levamisole treatment of wild type and *unc-119* animals causes hypercontraction of all body wall muscles and a strong dumpy appearance.

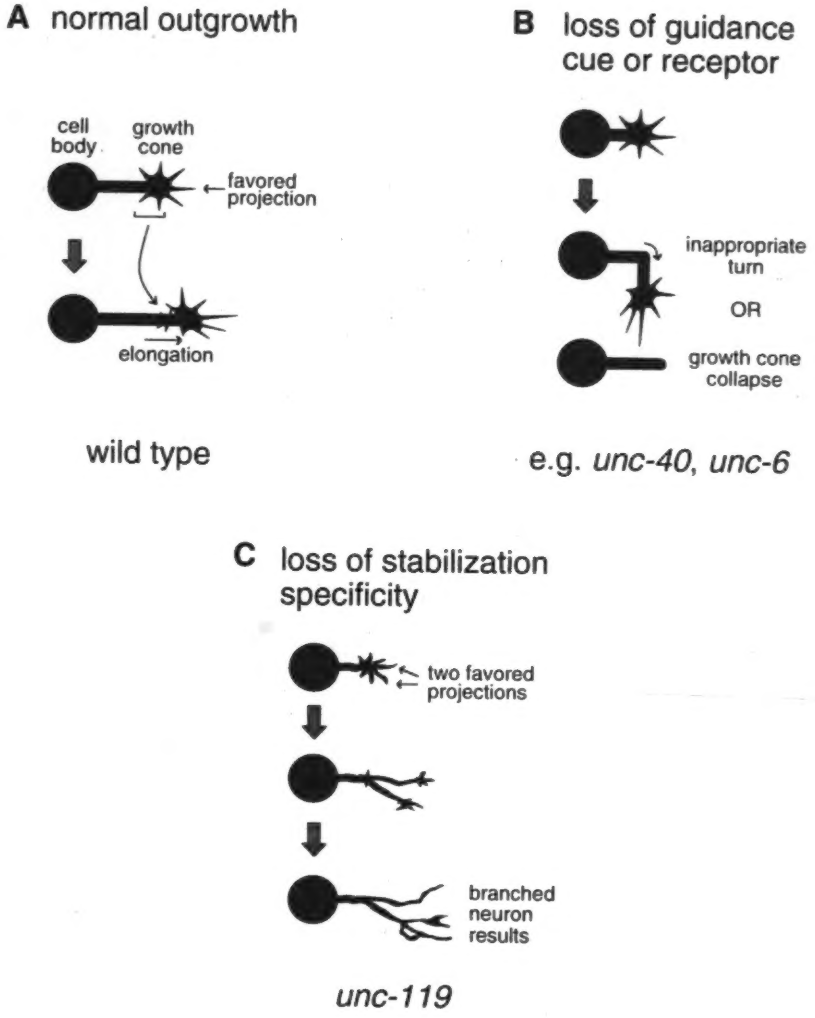


Figure 6.2 A possible model for UNC-119 function. **A**, Filopodial projections continually protrude and retract. Growth cone migration occurs when a filopodium is stabilized (e.g. because it is closer to a chemoattractant). Other filopodia collapse and an axoneme is formed as the growth cone moves forward. **B**, Mutations that disrupt the source of the signal might lead to inappropriate turns or growth cone collapse, but branched neurons are generally not seen. **C**, In *unc-119*, a branched neuron might result from disruption of the mechanism which allows a single filopodium to be stabilized.

6.12 Bibliography

Albertson, D. G., and Thomson, J. N. (1976). The pharynx of *Caenorhabditis elegans*. Phil. Trans. R. Soc. Lond.B Biol. Sci. *275*, 299-325.

Antebi, A., Norris, C. R., Hedgecock, E. M., and Garriga, G. (1996). Cell and Growth Cone Migrations. in *C. elegans II*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Avery, L., and Thomas, J. H. (1997). Feeding and Defecation. in *C. elegans* II. Cold Spring Harbor Laboratory, New York.

Chan, S. S. -Y., Zheng, H., Su, M. -W., Wilk, R., Killeen, M. T., Hedgecock, E. M., and Culotti, J. G. (1996). UNC-40, a *C. elegans* homolog of DCC (Deleted in Colorectal Cancer), is required in motile cells responding to UNC-6 netrin cues. Cell *87*, 187-195.

Coburn, C. M., and Bargmann, C. I. (1996). A Putative Cyclic Nucleotide-Gated Channel Is Required for Sensory Development and Function in *C. elegans*. Neuron *17*, 695-706.

Croll, N. A. (1975). Components and patterns in the behavior of the nematode *Caenorhabditis elegans*. J. Zoology *176*, 159-176.

Croll, N. A., and Smith, J. M. (1978). Integrated behavior in the feeding phase of *Caenorhabditis elegans* (Nematoda). J. Zoology *184*, 507-517.

De Stasio, E., Lephoto, C., Azuma, L., Holst, C., Stanislaus, D., and Uttam, J. Characterization of Revertants of *unc-93(e1500)* in *Caenorhabditis elegans* Induced by *N*-ethyl-*N*-nitrosourea. Genetics *147*, 597-608.

Driscoll, M. and Kaplan, J. (1997). Mechanotransduction. in *C. elegans II*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Florio, S. K., Prusti, R. K., and Beavo, J. A. (1996). Solubilization of Membrane-bound Rod Phosphodiesterase by the Rod Phosphodiesterase Recombinant ∂ Subunit. J. Biol. Chem. *271*, 24036-24047.

Goshima, Y., Nakamura, F., Strittmatter, P., and Strittmatter, S. M. (1995). Collapsin-induced growth cone collapse mediated by an intracellular protein related to UNC-33. Nature *376*, 509-514.

Hedgecock, E. M., Culotti, J. G., Thomson, J. N., and Perkins, L. A. (1985). Axonal guidance mutants of *Caenorhabditis elegans* identified by filling sensory neurons with fluorescein dyes. Dev. Biol. *111*, 158-170.

Herman, R. (1997). Mosaic Analysis. in *Caenorhabditis elegans*: *Modern Biological Analysis of an Organism* (eds Epstein, H. F. and Shakes, D. C.) Methods Cell Biol. *48*, 123-146.

Higashide, T., Murakami, A., McLaren, M. J., and Inana, G. (1996). Cloning of the cDNA for a Novel Photoreceptor Protein. J. Biol. Chem. *271*, 1797-1804.

Horvitz, H.R., Chalfie, M., Trent, C., Sulston, J. E., and Evans, P. D. (1982). Serotonin and octopamine in the nematode *Caenorhabditis elegans*. Science *216*, 1012-1014.

Ishii, N., Wadsworth, W. G., Stern, B. D., Culotti, J. G., and Hedgecock, E. M. (1992). UNC-6, a laminin-related protein, guides cell and pioneer axon migrations in *C. elegans*. Neuron *9*, 873-881.

Kennedy, T. E., Serafini, T., de la Torre, J. R., and Tessier-Lavigne, M. (1994). Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord. Cell *78*, 425-435.

Kimble, J. E., and White, J. G. (1981). On the control of germ cell development in *Caenorhabditis elegans*. Dev. Biol. *81*, 208-219.

Koelle, M. R., and Horvitz, H. R. (1996). ELG-10 Regulates G Protein Signaling in the C. elegans Nervous System and Shares a Conserved Domain with Many Mammalian Proteins. Cell *84*, 115-125.

Komatsu, H., Mori, I., Rhee, J.-S., Akaike, N., and Ohshima, Y. (1996). Mutations in a Cyclic Nucleotide-Gated Channel Lead to Abnormal Thermosensation and Chemosensation in *C. elegans*. Neuron *17*, 707-718.

Leung-Hagesteijn, C., Spence, A. M., Stern, B. D., Zhou, Y., Su, M. -W., Hedgecock, E. M., and Culotti, J. G. (1992). UNC-5, a transmembrane protein with immunoglobulin and thrombospondin type I domains, guides cell and pioneer axon migrations in *C. elegans*. Cell *71*, 289-299.

Li, W., Herman, R. K., and Shaw, J. E. (1992). Analysis of the *Caenorhabditis elegans* axonal guidance and outgrowth gene *unc-33*. Genetics *132*, 675-689.

Mello, C. and Fire, A. (1997). DNA Transformation. in *Caenorhabditis elegans*: *Modern Biological Analysis of an Organism* (eds Epstein, H. F. and Shakes, D. C.) Methods Cell Biol. *48*, 451-482.

Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. EMBO J. *10*, 3959-3970.

Moerman, D. G., Plurad, S., Waterston, R. H., and Baillie, D. L. (1982). Mutations in the *unc-54* myosin heavy chain gene of *Caenorhabditis elegans* that contractility but not muscle structure. Cell *29*, 773-781.

Murray, J. W., Edmonds, B. T., Liu, G., and Condeelis, J. (1996). Bundling of Actin Filaments by Elongation Factor 1 α Inhibits Polymerization at Filament Ends. J. Cell Biol. *135*, 1309-1321.

Nicholls, J. G., Martin, A. R., and Wallace, B. G. (eds.) (1992). *From neuron to brain: a cellular and molecular approach to the function of the nervous system*, 3rd ed. Sinauer Associates, Inc. Sunderland, Massachusetts.

Ogura, K., Wicky, C., Magnenat, L., Tobler, H., Mori, I., Muller, F., and Ohshima, Y. (1994). *Caenorhabditis elegans unc-51* gene required for axonal elongation encodes a novel serine/threonine kinase. Genes Dev. *84*, 2389-2400.

Ségalat, L., Elkes, D. A., and Kaplan, J. M. (1995). Modulation of serotonin-controlled behaviors by G_0 in *Caenorhabditis elegans*. Science *267*, 1648-1651.

Svendsen, P. C., and McGhee, J. D. (1995). The *C. elegans* neuronally expressed homeobox gene ceh-10 is closely related to genes expressed in the vertebrate eye. Development *121*, 1253-1262

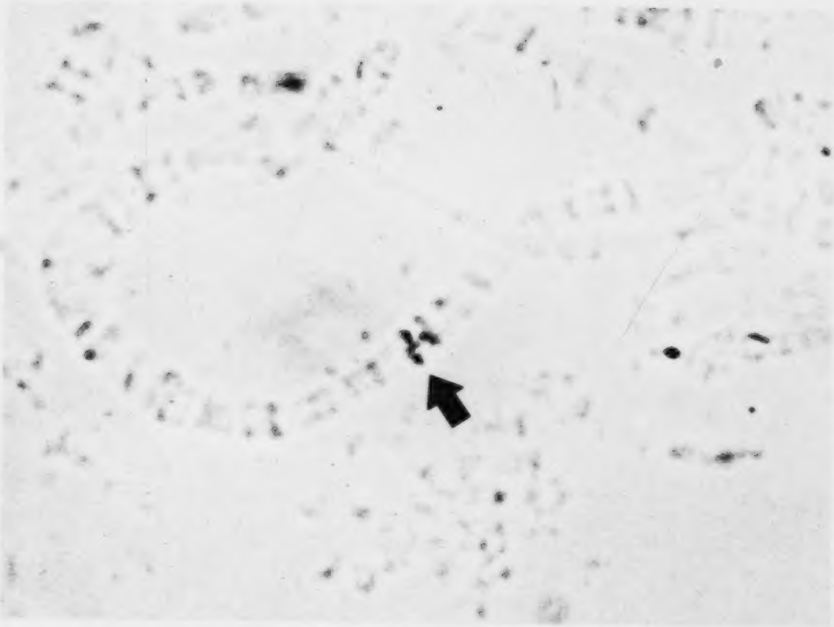
Tanaka, E. and Sabry, J. (1995). Making the Connection: Cytoskeletal Rearrangements during Growth Cone Guidance. Cell *83*, 171-176.

Wadsworth, W. G., Bhatt, H., and Hedgecock, E. M. (1996). Neuroglia and Pioneer Neurons Express UNC-6 to Provide Global and Local Netrin Cues for Guiding Migrations in C. elegans. Neuron *16*, 35-46.

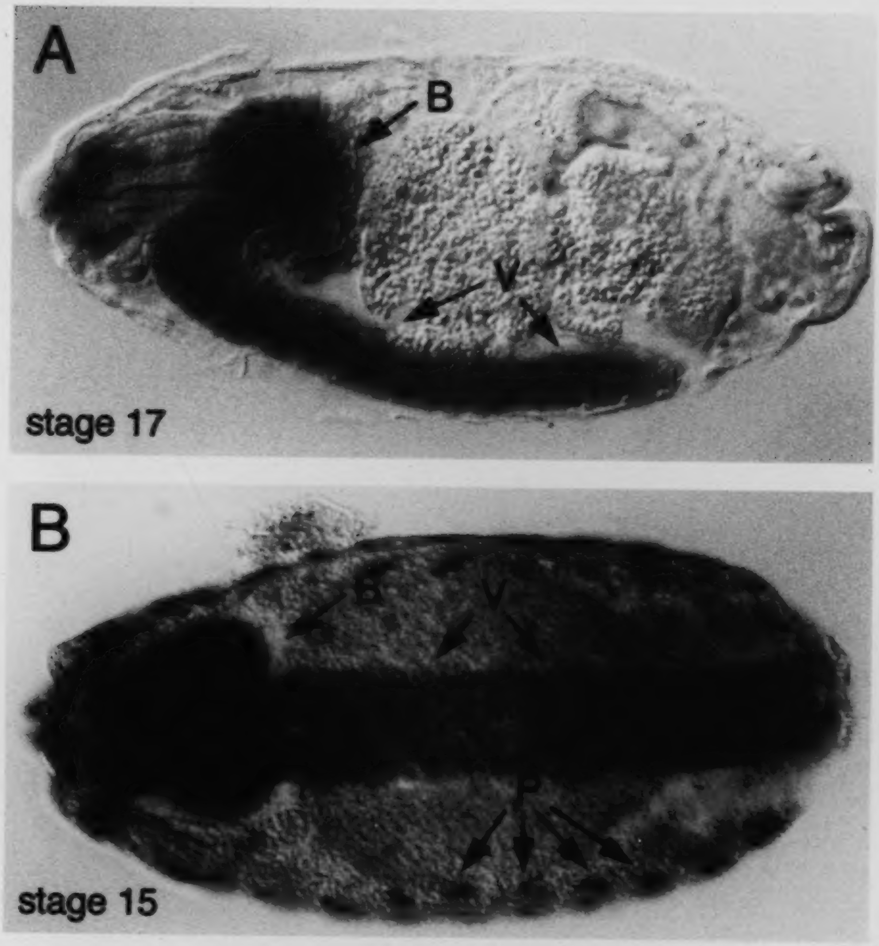
White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1986). The structure of the nervous system of the nematode *C. elegans.* Phil. Trans. R. Soc. Lond. B Biol. Sci. *314*, 1-340.

White, J., Southgate, E., and Thomson, J. N. (1992). Mutations in the *Caenorhabditis elegans unc-4* gene alter the syanptic input to ventral cord motor neurons. Nature *355*, 838-841.

Williamson, T., Gordon-Weeks, P. R., Schachner, M., and Taylor, J. (1996). Microtubule reorganization is obligatory for growth cone turning. PNAS (USA) *93*, 15221-15226.



Appendix 1 Localization of DmUNC119 to the X chromosome of *Drosophila melanogaster*, band 7A1, by *in situ* hybridization. This work was performed by D. Nash and E. Woloshyn, University of Alberta.



Appendix 2 Detection of DmUNC119 mRNA in *Drosophila* embryos by *in situ* hybridization. **A**, Stage 17 embryo, lateral view, showing expression in ventral nerve cord (V) and brain (B). **B**, Stage 15 embryo, ventral view. Expression occurs in ventral nerve cord (V), brain (B) and the peripheral nervous system (P). Anterior is to the left in both panels. This work was performed by Roger Jacobs, McMaster University, and Michael Gordon, McMaster University.